FIDUCIAL PROBABILITY THEORY
FOR DISTRIBUTIONS WITH A GROUP STRUCTURE

Rajinder Bir Hora*

Technical Report No. 33

*This research was supported by the National Science Foundation under Grant Number G19126.

Chapter 1. Introduction.

1.0. Summary

The subject of fiducial probability was introduced by Fisher about thirty years ago. Among the problems considered in this area have been those of fiducial estimation (Pitman (1939)) and fiducial prediction (Fisher (1935, 1956) and Ramsey and Buehler (1963)). This thesis, broadly speaking, is yet another small contribution in this direction.

It deals mainly with the problems of fiducial estimation and prediction for families of distributions with a group structure. Expectation identities useful for estimation and prediction purposes have been obtained. By application of these identities certain "best" estimators and predictors have been derived for certain location and scale parameter families. The expectation identity useful in estimation has also been applied to the estimation of an angular parameter in rotation families. Finally, some theorems (which correspond to theorems on invariant functions on the sample space given by Lehmann (1959)), concerning invariant functions on the sample and parameter spaces have been obtained.

1.1. Review of Previous Work.

Segal (1938) has derived fiducial distributions of several parameters. He begins by showing the existence of pivotals, then proceeds to suggest a method for obtaining the ancillaries and finally, by use of the pivotals, derives the fiducial density of several parameters. In Chapter 1, a rather simplified account for showing the existence of pivotals is given. Besides, the relation of standard pivotals to the more commonly known pivotals and ancillaries in case of location and scale families is considered.

Pitman (1939) considered the problem of fiducial estimation. He dealt with families of distributions having (i) one location, (ii) one location and one scale, and (iii) two location and one scale parameter, and obtained certain "best" estimators. In Chapter 2, an expectation identity for families of distributions with a group structure has been obtained for certain invariant functions, and Pitman's results are seen to be special cases of this identity.

Fisher was possibly first to use the term "fiducial prediction" in 1956. For a review of the work done in this area, the reader is referred to Buehler (1963), pages 22-23.

More recently Ramsey and Buehler (1963) have considered the problem of predicting a future observation (i.e., the $(n + 1)^{th}$ observation, if a sample of size n has been observed) in the case of one location and of one location and one scale parameter families. Proceeding on lines somewhat similar to those of Pitman (1939), certain "best" predictors were obtained. In Chapter 4, an expectation identity for a family of distributions with a group structure has been obtained. The identity can be applied to obtain "best" predictors, and results of Ramsey and Buehler (1963) are seen to be special cases of this identity.

In deriving these identities an extensive use of Fraser (1961), has been made since the formulation and the underlying assumptions are nearly the same. In Fraser (1961), a mathematical model involving transformation group is presented and the derivation of fiducial distributions and other related questions such as combining fiducial distributions have been considered.

Lehmann (1959) in Chapter 6, has considered invariant functions on a sample space. He has obtained theorems concerning: (i) relationship of the maximal invariant with an invariant function, (ii) a method for obtaining maximal invariant, (iii) the manner in which the parameter space can be shrunk by use of a maximal invariant on the parameter space. In Chapter 5 theorems corresponding to (i) and (ii) above have been obtained for invariant functions on sample and parameter spaces.

1.2. Contents of the Thesis.

This thesis, which is in five chapters, generally speaking considers (i) some general aspects of fiducial probability theory, (ii) the problems of fiducial estimation and prediction, and (iii) invariant functions on the sample and parameter spaces.

Chapter 1 begins with a review of the past work done in fiducial probability theory as related to the above problems. Papers of Segal (1938), Pitman (1939), Fraser (1961), and Ramsey and Buehler (1963) have been discussed. A simplified account on the existence of pivotals has been given in Section 1.3. It may be remarked that this topic has been dealt with by Segal (1938). In Section 1.4, the relation between the standard pivotals to the commonly known pivotals and ancillaries in location and scale families has been obtained.

In Chapter 2, the problem of estimation for families of distributions with a group structure is considered. The formulation of the problem is as follows:

Essentially, the following assumptions (which are nearly the same as that of Fraser (1961)) are made: (1) $(\underline{\mathrm{X}}, \beta_{\underline{\mathrm{X}}}, p^{\omega})$ is a probability space where $\omega\epsilon\Omega$, and Ω is the parameter space. (ii) There exist spaces $\mathcal{T}$ and $\mathcal{A}$ such that $\underline{\mathrm{X}}$ is in one-to-one correspondence with $\mathcal{T} \times \mathcal{A}$. (iii) There is a group $\mathcal{G}$ of measurable transformations on the sample space $\underline{\mathrm{X}}$ on to itself. (iv) The class of measures p^{ω} for $\omega\epsilon\Omega$ is closed under $\mathcal{G}$, i.e., for $X\epsilon\beta_{\underline{\mathrm{X}}}$, there exists a $g^{*}\omega\epsilon\Omega$ such that $p^{\omega}(X) = p^{g^{*}\omega}(gX)$. (v) The induced group $\mathcal{G}^{*}$ of transformations on the parameter space Ω is exactly transitive. (vi) There exists a

Haar measure on space $\mathcal{G}$ and thereby on several other spaces which are isomorphic to it.

Let E_R denote the conditional expectation given the ancillary and let E_f denote expectation with respect to the fiducial density of ω given x. Then, for any function $H(x, \omega)$ which satisfies the invariance condition $H(x, \omega) = H(gx, g^*\omega)$ it is shown that $E_R\ (H(x, \omega)) = E_f\ (H(x, \omega))$. It is seen (a counterexample is given) that the above condition on $H(x, \omega)$ is not necessary for the identity to hold. From this rather general and fundamental theorem on fiducial estimation, certain other theorems follow as special cases. In Sections 2.3 to 2.6, the problem of estimation in case of the following families of distributions have been considered: (i) One location parameter. (ii) One location parameter and one scale parameter. (iii) Two location parameters and one scale parameter. (iv) Two location parameters and two scale parameters. For all these families certain "best" invariant estimators have been obtained. In Section 2.7., some remarks have been made concerning special cases not dealt with in the preceeding sections.

In Chapter 3, the expectation identity obtained in Chapter 2 has been applied to rotation families. In Section 3.1, the formulation of the problem is essentially as follows: consider $X = (X_1, X_2)$ to be a random variable with a known density which has no parameter of its own. Then by rotation of the axes a family of random variables is generated. The problem is to obtain an estimate of α, the angle through which the axes have been rotated. Thus we are concerned with distributions on a circle. It is pointed out that such distributions pose special problems, e.g. the usual additive property of expectation is not necessarily valid. In Section 3.2, a brief account of the verification of the assumptions made in Chapter 2 is given.

In Section 3.3, the general conditional and fiducial densities are derived and examples concerning special cases of these are given. In Section 3.4, invariant estimators are characterized when the estimator is based on one or more bivariate observations. In Section 3.5, the definitions of expectation with respect to the conditional and fiducial densities are made. Also, the concept of symmetry for distributions on the circle is defined. Section 3.6, which is perhaps the basic section in this chapter, probability theory for distributions on a circle is considered. The terms quasi-mean and quasi-median are defined and certain lemmas concerning them have been obtained.

By use of fiducial theory certain "best" estimators have been obtained. Finally in this section the equivalence of fiducial and posterior distributions when prior distribution is taken to be uniform over the interval $(0, 2\pi)$ is discussed. In Section 3.7 "best" estimators for certain special cases have been obtained. In Section 3.8 a theorem for symmetric unimodal circular distributions is obtained. It deals with the minimization of a certain integral. An illustration of the theorem has also been given. In Section 3.9, a theorem by which a Bayes' estimate of α can be obtained in case of one or more bivariate observations is given. Certain corollaries which are special cases of the theorem are also given.

In Chapter 4, the problem of prediction for families of distributions which a group structure is considered. The formulation of the problem is as follows:

Essentially, the following assumptions are made: (i) $(\underline{\overline{X}}', \beta_{\underline{\overline{X}}'}, P^{\omega})$ is a probability space, where $\omega\epsilon\Omega$ and Ω is the parameter space. Also $\underline{\overline{X}}' = \underline{\overline{X}} \times \underline{\overline{X}}^{*}$, where $\underline{\overline{X}}$ is the space of "past" observations and $\underline{\overline{X}}'$ is the space of "future" observations. (ii)There exist spaces $\mathcal{T}$, $\mathcal{A}$ and $\mathcal{A}^{*}$ such that $\underline{\overline{X}}'$ is in one-to-one correspondence with $\mathcal{T} \times \mathcal{A} \times \mathcal{A}^{*}$ and $\underline{\overline{X}}$ is in one-to-one correspondence with $\mathcal{T} \times \mathcal{A}$. (iii)There is a group $\mathcal{G}$ of measurable transformations on the sample space $\underline{\overline{X}}'$ onto itself. (iv)The class of measures P^{ω} for $\omega\epsilon\Omega$ is closed under $\mathcal{G}$, i.e., for $X'\epsilon\beta_{\underline{\overline{X}}'}$ there exists $\underset{-}{g}^{*}\omega\epsilon\Omega$ such that

$P^{\omega}(X') = P^{g^{*}\omega}(gX')$. (v)The induced group $\underset{-}{\mathcal{G}}^{*}$ of transformations on the parameter space Ω is exactly transitive. (vi)There exists a Haar measure on space $\mathcal{G}$ and thereby on several other spaces which are isomorphic to it.

In this framework a new definition is given of the joint fiducial distribution of the parameter ω and future observations x^{*} given "past" observations x. Let E_R denote the conditional expectation given the ancillary a (which is based on the "past" observations). Let E_f denote the expectation with respect to the joint fiducial distribution of ω and $\underset{-}{x}^{*}$ given x. Then for any function $H(x',\omega)$ which satisfies the invariance condition $H(x',\omega) = H(gx',\underset{-}{g}^{*}\omega)$, it is shown that $E_R(H(x',\omega)) = E_f(H(x',\omega))$. From this general theorem on fiducial prediction, certain other theorems follow as special cases. In Sections 4.3. to 4.6., the problem of prediction in case of the following families of distributions have been considered: (i)one location parameter, (ii)one location parameter and one scale parameter, (iii)two location parameters and one scale parameter, (iv)two location parameters and two scale parameters. For all these families certain "best" predictors have been obtained. In Section 4.7., some remarks concerning other special cases have been made.

In Chapter 5, we consider invariant functions on the sample and parameter spaces. Theorems concerning the following have been obtained: (i) relationship of a maximal invariant with an invariant function, (ii) a method for obtaining a maximal invariant through subgroups. Examples bearing on these theorems and also an example of a function which is invariant but not maximal invariant have been given.

Finally, we want to make a few miscellaneous remarks concerning numbering, etc. The numbering of lemmas, corollaries, theorems, definitions, and remarks is continuous within each chapter. Also the figure before the decimal point in their numbering is the number of the chapter and the figure after is the number within the chapter. The numbering of equations is continuous within each chapter. The numbering of sections is on the similar lines. Section 2.2. means the second section in the second chapter, and Section 2.2.1 means the first subsection of the second section in the second chapter.

1.3.* On the Existence of Pivotals and Ancillaries.

In this section we consider the question of existence of pivotals and ancillaries. The existence of pivotals has been discussed in Segal (1938). Here we give a slightly different discussion of this subject.

Definition 1.1.

A function $P(x, \omega)$ for $x \epsilon \underline{\bar{X}}$ (the sample space) and $\omega \epsilon \Omega$ (the parameter space) is said to be pivotal if its distribution does not depend on ω.

Example 1.1.

Let X be a normal variate with mean 0 and variance 1. Then $X-\theta$ is a pivotal for it is distributed normally with mean 0 and variance 1.

Example 1.2.

Let X_1 and X_2 be independent normal variates each with mean θ and variance 1. Then X_1-X_2 is an ancillary for it is a normal variate with mean 0 and variance 2.

Remark 1.1.

The pivotals and ancillaries play an exceedingly important role in fiducial probability theory. Their use in derivation of fiducial distributions has been discussed, for example, by Segal (1938),

*This section is not essential to the understanding of the later sections of the thesis. Accordingly, the reader could omit it if he chose to do so.

Pitman (1939), Owen (1948), Fisher (1956), Tukey (1957), Fraser (1961), Brillinger (1962), and Buehler (1963).

Consider a random variable $X = (X_1, \ldots, X_n)$ whose density depends on the parameter $\omega = (\omega_1, \ldots, \omega_r)$, $(r \leq n)$. Suppose a sufficient statistic $T = (T_1, \ldots, T_r)$ exists for the parameter ω. If $F(T_i | t_{i-1}, \ldots, t_1; \omega)$, $(i = 1, \ldots, r)$, are continuous, then each one of them is distributed uniformly over the interval $(0, 1)$. Hence they are pivotals. The other possibility is that no such T exists. Assume in that case that there exist n-r jointly ancillary statistics $g_i(x_1, \ldots, x_n)$, $i = 1, \ldots, n-r$, (whose joint distribution does not depend upon the parameter ω). If $F(X_i | x_{i-1}, \ldots, x_1, \omega)$, $(i = 1, \ldots, n)$, are continuous, then each of them is uniformly distributed over the interval $(0, 1)$. Hence they are pivotals. According to Segal (1938), it is possible to obtain functions $\psi_i (i = 1, \ldots, n-r)$ such that $g_i(x_1, \ldots, x_n) = \psi_i(u_1, \ldots, u_n)$ for $i = 1, \ldots, n-r$, where $u_i = F(X_i | x_{i-1}, \ldots, x_1; \omega)$ for $i = 1, \ldots, n$. We may then take any r functions of u_i's which are functionally independent of ψ_i's to be the pivotals. In this way the n standard pivotals u_i are transformed to a set of n-r ancillaries and r pivotals. The joint conditional distribution of the r pivotals given the n-r ancillaries may be used to obtain the fiducial distribution of ω given x.

Remark 1.2.

Though (as has been shown by Segal and indicated above) the pivotals u_i, $(i = 1, \ldots, n)$, always exists, the question of existence of ancillaries g_i, $(i = 1, \ldots, n-r)$, is yet an unsolved problem.

1.4.* Relationship of Pivotals and Ancillaries with the Segal Forms in the Case of Location and Scale Parameter Families.

In this section we consider for location and scale families the relationship of standard (Segal) forms of pivotals with the known forms of pivotals and the known forms of ancillaries. The known pivotals and ancillaries are expressed in terms of the standard pivotals.

We begin with the case of one location parameter. We have that $F(X_1, \theta)$ is a pivotal and also:

$$u_1 = F(x_1, \theta) = \int_{-\infty}^{x_1} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1-\theta, \ldots, x_n-\theta)\, dx_1 \ldots dx_n, \tag{1}$$

$$= \int_{-\infty}^{x_1-\theta} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f(x_1', \ldots, x_n')\, dx_1' \ldots dx_n',$$

where
$$x_i-\theta = x_i' \quad (i = 1, \ldots, n). \tag{2}$$

Thus $u_1 = G(x_1-\theta)$,

where $g(x_1') = \int_{-\infty}^{\infty} f(x_1', \ldots, x_n')\, dx_2' \ldots dx_n'$ and

$$G(x_1) = \int_{-\infty}^{x_1} g(x_1')\, dx_1'.$$

Therefore, we have from (2) that :

$$x_1-\theta = G^{-1}(u_1). \tag{3}$$

Thus the more commonly known pivotal $x_1-\theta$ is a function of the standard pivotal. Furthermore, if $F(x_1^*, \theta) \geq F(x_1^{**}, \theta)$, then $x_1^* \geq x_1^{**}$ or

* Sections marked * are not essential to the understanding of the later sections of the thesis.

equivalently $x_1^{*}-\theta \geqq x_1^{**}-\theta$. It follows that $x_1-\theta$ is a monotone function of the standard pivotal. We also have that:

$$u_2 = F(x_2|x_1,\theta) = \frac{\int_{-\infty}^{x_2}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1-\theta,\, x_2-\theta,\, \ldots,\, x_n-\theta)\, dx_2\ldots dx_n}{\int_{-\infty}^{\infty} f(x_1-\theta,\, x_2-\theta,\, \ldots,\, x_n-\theta)\, dx_2\ldots dx_n} \tag{4}$$

$$= \frac{\int_{-\infty}^{x_2-\theta}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(x_1-\theta,\, x_2',\, \ldots,\, x_n')\, dx_2'\ldots dx_n'}{\int_{-\infty}^{\infty} f(x_1-\theta,\, x_2',\, \ldots,\, x_n')\, dx_2'\ldots dx_n'} \tag{5}$$

Thus we have that:

$$u_2 = H(x_2-\theta,\, x_1-\theta), \tag{6}$$

where $h(x_2',\, x_1-\theta) = \dfrac{\int_{-\infty}^{\infty} f(x_1-\theta,\, x_2',\, \ldots,\, x_n')\, dx_3'\ldots dx_n'}{\int_{-\infty}^{\infty} f(x_1-\theta,\, x_2',\, \ldots,\, x_n')\, dx_2'\ldots dx_n'}$

and $\quad H(x_2,\, x_1-\theta) = \int_{-\infty}^{x_2} h(x_2',\, x_1-\theta)\, dx_2'$.

Finally we have that:

$$x_2-x_1 = (x_2-\theta) - (x_1-\theta) = H_1^{-1}\,(u_2,\, G^{-1}(u_1)) - G^{-1}(u_1), \tag{7}$$

where H_1^{-1} denotes the inverse of H with respect to the first argument and G^{-1} denotes the inverse of G.

Thus X_2-X_1, which is the more commonly known ancillarly, can be expressed as a function of the first two pivotals. Similarly x_i-x_1

$(i = 3, \ldots, n)$ is also expressible as a function of the standard pivotals. We thus obtain one pivotal and $(n-1)$ ancillaries and accordingly the n standard pivotals are transformed into 1 pivotal and $(n-1)$ ancillaries.

Next, we consider the case of one location parameter and one scale parameter. We have that $F(x_1, \theta, \sigma)$ is a pivotal and also:

$$t_1 = F(x_1, \theta, \sigma) = \int_{-\infty}^{x_1} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \sigma^{-n} f(\frac{x_1-\theta}{\sigma}, \ldots, \frac{x_n-\theta}{\sigma}) \, dx_1 \ldots dx_n. \tag{8}$$

$$= \int_{-\infty}^{\frac{x_1-\theta}{\sigma}} \int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} f(x_1', \ldots, x_n') \, dx_1' \ldots dx_n',$$

$$\text{where } \frac{x_i-\theta}{\sigma} = x_i' \quad (i = 1, \ldots, n) \tag{9}$$

Thus,

$$t_1 = G_1(\frac{x_1-\theta}{\sigma}), \tag{10}$$

where, $g_1(x_1') = \int_{-\infty}^{\infty} f(x_1', \ldots, x_n') \, dx_2' \ldots dx_n'$.

and, $G_1(x_1) = \int_{-\infty}^{x_1} g_1(x_1') \, dx_1'$.

Therefore, we have that:

$$\frac{x_1-\theta}{\sigma} = G_1^{-1}(t_1). \tag{11}$$

Hence, the commonly known pivotal is a function of the standard pivotal. Furthermore, by an argument similar to that used in the case of one location parameter it can be shown that $\frac{X_1-\theta}{\sigma}$ is a monotone function of the standard pivotal $F(X_1, \theta, \sigma)$. By rather similar caculations a second pivotal, e.g. $\frac{x_2-\theta}{\sigma}$ or $\frac{x_2-x_1}{\sigma}$ can also be obtained as a function of the standard pivotals.

We also have that:

$$t_2 = F(x_2|x_1, \theta, \sigma) = \frac{\int_{-\infty}^{x_2}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}\sigma^{-n} f(\frac{x_1-\theta}{\sigma}, \frac{x_2-\theta}{\sigma}, \dots, \frac{x_n-\theta}{\sigma})dx_2\dots dx_n}{\int_{-\infty}^{\infty}\sigma^{-n} f(\frac{x_1-\theta}{\sigma}, \frac{x_2-\theta}{\sigma}, \dots, \frac{x_n-\theta}{\sigma})\, dx_2\dots dx_n} \qquad (12)$$

$$= \frac{\int_{-\infty}^{\frac{x_2-\theta}{\sigma}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(\frac{x_1-\theta}{\sigma}, x_2', \dots, x_n')\, dx_2'\dots dx_n'}{\int_{-\infty}^{\infty} f(\frac{x_1-\theta}{\sigma}, x_2', \dots, x_n')\, dx_2'\dots dx_n'}, \qquad (13)$$

where $\frac{x_r-\theta}{\sigma} = x_r'$ $(1 = 2, \dots, n)$

Thus we have that:

$$t_2 = G_2(\frac{x_2-\theta}{\sigma}, \frac{x_1-\theta}{\sigma}), \qquad (14)$$

where $$g_2(x_2', \frac{x_1-\theta}{\sigma}) = \frac{\int_{-\infty}^{\infty} f(\frac{x_1-\theta}{\sigma}, x_2', \dots, x_n')\, dx_3'\dots dx_n'}{\int_{-\infty}^{\infty} f(\frac{x_1-\theta}{\sigma}, x_2', \dots, x_n')\, dx_2'\dots dx_n'}.$$

and $$G_2(x_2, \frac{x_1-\theta}{\sigma}) = \int_{-\infty}^{x_2} g_2(x_2', \frac{x_1-\theta}{\sigma})\, dx_2'.$$

Furthermore,

$$t_3 = F(x_3 | x_2, x_1, \theta, \sigma)$$

$$= \frac{\int_{-\infty}^{x_3}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} \sigma^{-n} f(\frac{x_1-\theta}{\sigma}, \frac{x_2-\theta}{\sigma}, \frac{x_3-\theta}{\sigma}, \ldots, \frac{x_n-\theta}{\sigma})\, dx_3 \ldots dx_n}{\int_{-\infty}^{\infty} \sigma^{-n} f(\frac{x_1-\theta}{\sigma}, \frac{x_2-\theta}{\sigma}, \frac{x_3-\theta}{\sigma}, \ldots, \frac{x_n-\theta}{\sigma})\, dx_3 \ldots dx_n}, \tag{15}$$

$$= \frac{\int_{-\infty}^{\frac{x_2-\theta}{\sigma}}\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty} f(\frac{x_1-\theta}{\sigma}, \frac{x_2-\theta}{\sigma}, x_3', \ldots, x_n')\, dx_3' \ldots dx_n'}{\int_{-\infty}^{\infty} f(\frac{x_1-\theta}{\sigma}, \frac{x_2-\theta}{\sigma}, x_3', \ldots, x_n')\, dx_3' \ldots dx_n'}, \tag{16}$$

where $\frac{x_i-\theta}{\sigma} = x_i'$ $(i = 3, \ldots, n)$.

Thus we have that:

$$t_3 = G_3\,(\frac{x_3-\theta}{\sigma}, \frac{x_2-\theta}{\sigma}, \frac{x_1-\theta}{\sigma}), \tag{17}$$

where $g_3(x_3', \frac{x_2-\theta}{\sigma}, \frac{x_1-\theta}{\sigma}) = \dfrac{\int_{-\infty}^{\infty} f(\frac{x_1-\theta}{\sigma}, \frac{x_2-\theta}{\sigma}, x_3', \ldots, x_n')\, dx_4' \ldots dx_n'}{\int_{-\infty}^{\infty} f(\frac{x_1-\theta}{\sigma}, \frac{x_2-\theta}{\sigma}, x_3', \ldots, x_n')\, dx_3' \ldots dx_n'}$,

and $\quad G_3(x_3, \frac{x_2-\theta}{\sigma}, \frac{x_1-\theta}{\sigma}) = \int_{-\infty}^{x_3} g_3(x_3', \frac{x_2-\theta}{\sigma}, \frac{x_1-\theta}{\sigma})\, dx_3'$.

Therefore, $\frac{x_3-x_1}{x_2-x_1}$, which is an ancillary can be expressed in terms of

t_1, t_2, and t_3 as follows:

$$\frac{x_3-x_1}{x_2-x_1} = \frac{\frac{x_3-\theta}{\sigma} - \frac{x_2-\theta}{\sigma}}{\frac{x_2-\theta}{\sigma} - \frac{x_1-\theta}{\sigma}} = \frac{G_3^{-1}(t_3, G_2^{-1}(t_2, G_1^{-1}(t_1)), G_1^{-1}(t_1)) - G_2^{-1}(t_2, G_1^{-1}(t_1))}{G_2^{-1}(t_2, G_1^{-1}(t_1)) - G_1^{-1}(t_1)}. \tag{18}$$

Also, (n-3) other ancillaries, e.g. $\frac{x_i-x_1}{x_2-x_1}$ (i = 4, ..., n) or $\frac{x_i-x_2}{x_2-x_1}$ (i = 4, ...n) can be obtained as functions of the ω standard pivotals. We thus can obtain two pivotals and (n-2) ancillaries as functions of n standard pivotals as was claimed by Segal.

The method suggested here for obtaining pivotals and ancillaries can easily be applied to more general cases, e.g. the case of families of distributions with two location parameters and two scale parameters. One could obtain in this case the four pivotals $\frac{x_1-\theta}{\sigma_1}$, $\frac{x_2-\theta_1}{\sigma_1}$; $\frac{x_1-\theta_2}{\sigma_2}$; $\frac{y_2-\theta_2}{\sigma_2}$ and the n + m - 4 ancillaries $\frac{x_{i+2}-x_1}{x_2-x_1}$ (i = 1, ..., m-2) and $\frac{y_{j+2}-y_1}{y_2-y_1}$ (j = 1, ..., n-2).

Chapter 2. Estimation.

2.0. Introduction.

The problem of estimation for families of distributions with location and scale parameters has been considered by Pitman (1939). This chapter considers a more general problem, namely the problem of estimation for families of distributions with a group structure. Pitman's problem is seen to be a special case of this.

2.1. Formulation of the Problem.

The formulation of the problem is as follows. We make the following assumptions.

Assumption 1.

Let $(\underline{\bar{X}}, \beta_{\underline{\bar{X}}}, P^{\omega})$ be a probability space where $\omega \epsilon \Omega$ and Ω is the parameter space.

Assumption 2.

There exist spaces $\mathcal{T}$ and $\mathcal{A}$ such that $\underline{\bar{X}}$ is in one-to-one correspondence with $\mathcal{T} \times \mathcal{A}$. For elements x,t,a of these spaces this correspondence will be denoted by:

$$x \longleftrightarrow (t,a) \tag{1}$$

Assumption 3.

Let $(\mathcal{T}, \beta_{\mathcal{T}})$, $(\mathcal{A}, \beta_{\mathcal{A}})$ and (Ω, β_{Ω}) be measurable spaces and $\beta_{\mathcal{T} \times \mathcal{A}}$ be the minimal σ-field which contains the cartesian product of the sets in the given σ-fields. Also assume that for $T \epsilon \beta_{\mathcal{T}}$ and $A \epsilon \beta_{\mathcal{A}}$, if

$$X = \{x: x \longleftrightarrow (t,a) \text{ for } t \epsilon T \text{ and } a \epsilon A\}, \tag{2(a)}$$

then $X \epsilon \beta_{\underline{\bar{X}}}$.

Finally assume that for any $X \epsilon \beta_{\underline{\overline{X}}}$, we have that if,

$$B = \{(t,a) <--> x \text{ for } x \epsilon X\}, \qquad (2(b))$$

then $B \epsilon \beta_{\mathcal{T} \times \mathcal{A}}$.

Since $x <--> (t,a)$, t is conditionally sufficient for ω given a.
Also, the measures P^{ω} on $\underline{\overline{X}}$ impose corresponding measures on $\mathcal{T} \times \mathcal{A}$
i.e. for every $T \epsilon \beta_{\mathcal{T}}$ and $A \epsilon \beta_{\mathcal{A}}$, $P^{\omega}(T \times A) = P^{\omega}(X)$, (3)
where X and $T \times A$ are related through (2(a)).

Remark 2.1.

There is no danger of confusion by using the same symbol P^{ω} for measures on spaces $\underline{\overline{X}}$ and $\mathcal{T} \times \mathcal{A}$.
Thus it is possible to index the conditional measures on $\mathcal{T}$ by $\omega \epsilon \Omega$.

Assumption 4.

There is a group $\mathcal{G} = \{g\}$ of (1-1) measurable transformations on the sample space $\underline{\overline{X}}$ on to itself and $(\mathcal{G}, \beta_{\mathcal{G}})$ is a measurable space. Also, assume there exists a left Haar measure μ on the space which has the invariance property given by:

$$\mu(gG) = \mu(G) \quad \text{for all} \quad g \epsilon \mathcal{G}, \quad \text{and} \quad G \epsilon \beta_{\mathcal{G}} . \qquad (4)$$

Assumption 5.

The class of measures P^{ω} for $\omega \epsilon \Omega$ is closed under $\mathcal{G}$.
Thus for any $g \epsilon \mathcal{G}$ and $\omega \epsilon \Omega$, there is a $\omega_g \epsilon \Omega$ such that for all $X \epsilon \beta_{\underline{\overline{X}}}$,

$$P^{\omega}(X) = P^{g^{*}\omega}(gX), \qquad (5)$$

where g^{*} is the function on Ω to Ω defined by $g^{*}\omega = \omega_g$.

Assumption 6.

For any $\omega_1, \omega_2 \in \Omega$, there is a single $g \in \mathcal{G}$ such that $g^*\omega_1 = \omega_2$. In other words, the mapping $g \longleftrightarrow g^*$ is an isomorphism and the group $\mathcal{G}^*$ is exactly transitive on Ω (i.e. for any ω_1, ω_2 there is a single g^* in $\mathcal{G}^*$ carrying ω_1 into ω_2). It may be remembered that in general (without Assumption 6) $g \longleftrightarrow g^*$ is an homorphism and not on an isomorphism.

Assumption 7.

For any $g \in \mathcal{G}$ and $X \in \overline{X}$, if $x \longleftrightarrow (t,a)$ then:

$$gx \longleftrightarrow (t_g, a) , \tag{6}$$

where t_g depends only on t and g and not on a.

Definition 2.1.

Let g' be defined by:

$$g't = t_g. \tag{7}$$

It is seen that $\mathcal{G}' = \{g'\}$ is a group that is isomorphic to $\mathcal{G}$.

Assumption 8.

$\mathcal{G}'$ is exactly transitive on $\mathcal{T}$.

Representation of spaces $\mathcal{T}$ and Ω in terms of $\mathcal{G}'$ and $\mathcal{G}^*$.

Let x_0 and ω_0 be arbitrary but fixed reference points in the sample space $\overline{X}$ and parameter space Ω. If $x_0 \longleftrightarrow (t_0, a_0)$, then t_0 and a_0 are taken to be the corresponding reference points in spaces $\mathcal{T}$ and $\mathcal{A}$.

Let g'_t be the unique transformation in $\mathcal{G}'$ which for each a, carries t_0 into t and g'_ω be the transformation which carries the conditional variable t_0 with ω_0 distribution into a conditional variable t with ω distribution. Also, let g^*_ω be the transformation that carries ω_0 in to ω. We will denote g'_t and g'_ω in $\mathcal{G}'$ by t and ω respectively, and g^*_ω in $\mathcal{G}^*$ by ω^*.

Remark 2.2.

The correspondence among the elements of groups of transformations $\mathcal{G}$, $\mathcal{G}'$, $\mathcal{G}^*$ is obvious.

Remark 2.3.

Since $\mathcal{G}'$ and $\mathcal{T}$ are spaces which are isomorphic to $\mathcal{G}$, the measure μ imposes corresponding measures on the spaces.

Assumption 9.

Assume that $\underline{\overline{X}} \subseteq R_n$ (the n-dimensional Euclidean space). Assume that for each ω, the measure P^ω is absolutely continuous with respect to the n-dimensional Lebesgue measure L_n (i.e., $P^\omega \ll L_n$).

Then by the Radon-Nikodym Theorem, there exists a L_n-measurable function $p(x,\omega)$ such that:

$$P^\omega(X) = \int_X p(x,\omega)\, dL_n(x), \tag{8}$$

for every $X \in \beta_{\underline{\overline{X}}}$.

Lemma 2.1.

If (5) holds, then for all g,ω:

$$p(x,\omega) = p(gx, g^*\omega) \quad \text{a.e. } L_n \tag{9}$$

where $g \longleftrightarrow g^*$.

Proof:

Since by (5) for $X \epsilon \beta_{\underline{X}}$, $P^{\omega}(X) = P^{g^* \omega}(gX)$, we have that:

$$\int_X p(x,\omega)dL_n(x) = \int_X p(gx,g^*\omega)dL_n(x). \tag{10}$$

But (10) is equivalent to (9). This establishes the lemma.

Lemma 2.2.

If for all x, g and ω (where $x \longleftrightarrow (t,a)$ and $g \longleftrightarrow g' \longleftrightarrow g^*$),

$$H(x,\omega) = H(gx,g^*\omega), \tag{11}$$

then $H(x,\omega)$ can be expressed in the form $H'(\omega^{-1}t,a)$.

Proof:

The correspondence $x \longleftrightarrow (t,a)$ defines a function H" according to

$$H(x,\omega) = H''(t,a,\omega).$$

Also,

$$H(gx,g^*\omega) = H''(g't,a,g^*\omega).$$

If we put $g^* = (\omega^*)^{-1}$ and hence $g' = \omega^{-1}$, then $H(x,\omega)$ is seen to be of the form $H'(w^{-1}t,a)$.

Assumption 10.

Assume that there exists a measure λ on $(\mathcal{A}, \beta_{\mathcal{A}})$ and a $\beta_{\mathcal{T} \times \mathcal{A}}$-measurable function h(t,a) such that every $S \epsilon B_n$ (the n-dimensional Borel field), we have that:

$$\int_S dL_n(x) = \int_{S'} h(t,a)d\mu(t)d\lambda(a) \quad , \tag{12}$$

where $S' = \{(t,a) : x \longleftrightarrow (t,a) \text{ if } x \epsilon S\}$.

Definition 2.2.

For $A \epsilon \beta_{\mathcal{A}}$, define:

$$P_2^{\omega}(A) = P^{\omega}(\mathcal{T} \times A). \tag{13}$$

Remark 2.4.

$P_2(\cdot)$ is a probability measure on $\mathcal{A}$.

Lemma 2.3.

$P_2^{\omega}(\cdot)$ does not depend on ω.

Proof:

Since by (3), (5) and $\mathcal{T}_g = \mathcal{T}$, where $\mathcal{T}_g = \{t_g : t \epsilon \mathcal{T}\}$, we have that:

$$P_2^{\omega}(A) = P^{\omega}(\mathcal{T} \times A) = P^{\omega}(X) = P^{g^*\omega}(gX)$$
$$= P^{g^*\omega}(\mathcal{T}_g \times A) = P^{g^*\omega}(\mathcal{T} \times A) = P_2^{g^*\omega}(A),$$

where X and $\mathcal{T} \times A$ are related through (2(a)).

But, $\{g^*\omega : g^* \epsilon \mathcal{G}^*\} = \{\omega : \omega \epsilon \Omega\}$. It follows that $P_2(\cdot)$ does not depend upon ω. Due to (13), we have that P^{ω} considered as a measure on space $\mathcal{A}$ for fixed T is absolutely continuous with respect to P_2. Hence we can define a $\beta_{\mathcal{A}}$ -measurable function $P_1^{\omega}(T \mid a)$ by use of Radon-Nikodyn Theorem according to:

$$P^{\omega}(T \times A) = \int_A P_1^{\omega}(T \mid a) dP_2(a). \tag{14}$$

Remark 2.5.

For almost all a, $P_1(\cdot \mid a)$ is a conditional probability measure on space $\mathcal{T}$.

Assumption 11.

Assume that for all a, $P_1 << \mu$.

By Assumption I1 and application of Radon-Nikodyn Theorem, we define a $\beta_{\mathcal{T}}$-measurable function $p_1(t \mid a,\omega)$ according to:

$$P_1^{\omega}(T \mid a) = \int_{t\in T} p_1(t \mid a,\omega)d\mu(t). \tag{15}$$

Remark 2.6.

$p_1(t \mid a,\omega)$ is a conditional density on space $\mathcal{T}$ with respect to the Haar measure μ.

Lemma 2.4.

$$p_1(t \mid a,\omega) = p_1'(\omega^{-1}t \mid a). \tag{16}$$

By (14), we have that for all $A\in\beta_{\mathcal{A}}$,

$$P^{\omega}(T\times A) = \int_A P_1^{\omega}(T\times A)dP_2(a).$$

and similarly, letting $T_g = \{t_g : t\in T\}$,

$$P^{g^*\omega}(T_g\times A) = \int_A P_1^{g^*\omega}(T_g \mid a)dP_2(a).$$

Since by an agrument similar to that used in Lemma 2.3.

$P^{\omega}(T\times A) = P^{g^*\omega}(T_g\times A)$ we have that for all $A\in\beta_{\mathcal{A}}$,

$$\int_A P_1^{\omega}(T \mid a)dP_2(a) = \int_A P_1^{g^*\omega}(T_g \mid a)dP_2(a).$$

Hence, $P_1^{\omega}(T \mid a) = P_1^{g^*\omega}(T_g \mid a)$, a.e. P_2.

Also by (15),

$$P_1^{\omega}(T \mid a) = \int_{t\in T} p_1(t \mid a,\omega)d\mu(t).$$

and similarly,

$$p_1^{g^*\omega}(T_g \mid a) = \int_{t_g \in T_g} p_1(t_g \mid a; g^*\omega)d\mu(t).$$

$$= \int_{t \in T} p_1(t_g \mid a, g^*\omega)d\mu(t). \quad \text{(by (4))}.$$

Hence,

$$p_1(t \mid a;\omega) = p_1(t_g \mid a, g^*\omega), a.e.\mu.$$

Put $g^* = \omega^{-1}$, and we have that (16) holds.

Definition 2.3.

Define a measure ν on the space $\mathcal{T}$ by:

$$\nu(T) = \mu(T^{-1}), \text{ for } T \in \beta_{\mathcal{T}}. \tag{17}$$

Remark 2.7.

ν is a right Haar measure.

Remark 2.8.

$$\mu(Tg') = \Delta(g')\mu(T), \text{ for } T \in \beta_{\mathcal{T}}, \tag{18}$$

where $\Delta(g')$ is the modular function. This is due to the fact that invariant measures are unique up to a constant and $\mu_{g'}$ where $\mu_{g'}(T) = \mu(Tg')$ is another left Haar measure on space $\mathcal{T}$.

Definition 2.4.

Following the lines of Fraser (1961), we define the fiducial density of ω given x with respect to the measure ν by:

$$p(\omega \mid x) = p_1(t \mid a,\omega)\Delta(t) = p_1'(\omega^{-1}t \mid a)\Delta(t), \tag{19}$$

where $\Delta(t)$ is the modular function.

2.2. Expectation Identity.

In this Section we obtain an expectation identity, $E_R(H(x,\omega)) = E_f(H(x,\omega))$, where $H(x,\omega)$ is a function for which (11) holds. It will be seen that (11) is a sufficient but not a necessary condition for the validity of the expectation identity. We begin with the theorem which proves the above mentioned expectation identity.

2.2.1. Theorem Concerning Expectation Identity.

Assume:

(1) Assumptions 1-11 are satisfied.

(2) $H(x,\omega)$ is an invariant function, i.e. (11) holds.

(3) E_R denotes the conditional expectation with respect to the conditional density $p_1(t \mid a,\omega)$ and E_f denotes the expectation with respect to the density(19).

Then:

$$E_R(H(x,\omega)) = E_f(H(x,\omega)). \tag{20}$$

Proof:

$$E_R(H(x,\omega)) = \int_{t\in T} H'(\omega^{-1}t,a)p_1(t \mid a,\omega)d\mu(t), \text{ by Lemma 2.2.}$$

$$= \int_{t\in T} H'(\omega^{-1}t,a)p_1'(\omega^{-1}t \mid a)d\mu(\omega^{-1}t), \text{ by (16) and (4).}$$

$$= \int_{s\in\omega^{-1}T=T} H'(s,a)p_1'(s \mid a)d\mu(s), \; s = \omega^{-1}t.$$

$$= \int_{\omega\in\Omega} H'(\omega^{-1}t,a)p_1'(\omega^{-1}t \mid a)d\mu(\omega^{-1}t), \text{ for a fixed } t.$$

$$= \int_{\omega\in\Omega} H'(\omega^{-1}t,a)p_1'(\omega^{-1}t \mid a)\Delta(t)d\mu(\omega^{-1}), \text{ by (18).}$$

$$= \int_{\omega\in\Omega} H'(\omega^{-1}t,a)p_1'(\omega^{-1}t \mid a)\Delta(t)d\nu(\omega), \text{ by (17).}$$

$$= E_f(H(x,\omega)), \text{ by (19).}$$

2.2.2. Counterexample.

The following is an example which shows that (11) is not a necessary condition for (20) to hold.

Consider the case when X is a uniform variate over the interval $(\theta,\theta + 1)$. Let $H(x,\theta)$ be given as follows:

$$\begin{aligned} H(x,\theta) &= -(x + \theta),\ \theta \leq x \leq \theta + \tfrac{1}{4} \text{ and } \theta + \tfrac{3}{4} < x \leq \theta + 1. \\ &= (x + \theta),\ \theta + \tfrac{1}{4} < x \leq \theta + \tfrac{3}{4} \\ &= 0, \quad \text{otherwise.} \end{aligned}$$

$$\text{Then, } E_R(H(x,\theta)) = \int_{\theta}^{\theta+\frac{1}{4}} -(x+\theta)dx + \int_{\theta+\frac{1}{4}}^{\theta+\frac{3}{4}} (x+\theta)dx + \int_{\theta+\frac{3}{4}}^{\theta+1} -(x+\theta)dx = 0$$

$$\text{Also, } E_f(H(x,\theta)) = \int_{x-\frac{1}{4}}^{x} -(x+\theta)d\theta + \int_{x-\frac{3}{4}}^{x-\frac{1}{4}} (x+\theta)d\theta + \int_{x-1}^{x-\frac{3}{4}} -(x+\theta)d\theta = 0.$$

Thus, $E_R(H(x,\theta)) = E_f(H(x,\theta))$ and yet $H(x,\theta)$ is not a function for which (11) holds.

Remark 2.9.

Of course, one can always construct functions $H(x,\omega)$ for which (11) holds a.e. and (20) still holds.

2.3. Case of One Location Parameter.

In this section we consider a theorem which deals with families of distributions with one location parameter θ.

Definition 2.5.

An estimator $\hat{\theta}(x_1, \ldots, x_n)$ of θ is said to be invariant if for any a, $-\infty < a < \infty$,

$$\hat{\theta}(x_1, \ldots, x_n) - a = \hat{\theta}(x_1 - a, \ldots, x_n - a). \tag{21}$$

Definition 2.6.

For $\underline{\bar{X}} = R_n$ define g on $\underline{\bar{X}}$ as follows:

$$gx = (x_1 - g, \ldots, x_n - g). \tag{22}$$

For $\Omega = R_1$, define g^* on $\Omega = \{\theta\}$ as follows:

$$g^*\theta = \theta - g. \tag{23}$$

Pitman (1939) obtained expectation identities for functions of the form $(\hat{\theta}(x_1, \ldots, x_n) - \theta)^m$ and $|\hat{\theta}(x_1, \ldots, x_n) - \theta|^m$ for $m \geq 0$ and for invariant estimators $\hat{\theta}$. With the above definition of g and g^* and with $\theta = \omega$ it is easily shown that these functions satisfy the invariance relation (11) for H. Thus the following theorem, which will be shown to follow from Theorem 2.1., is very similar to Pitman's result. It is actually slightly more general in two respects: the function H is more general; and the variates $x_1, \ldots, x_n$ are not necessarily independent or identically distributed.

Theorem 2.2.

Let E_R denote the conditional expectation over a region R in which the ancillaries $a_i = x_{i+1} - x_1 (i = 1, \ldots, n-1)$ are fixed and E_f denote the expectation with respect to the fiducial density of θ given x. Assume that:

(1) x has density with respect to the Lebesgue measure L_n given by:

$$p(x,\theta) = f(x_1 - \theta, \ldots, x_n - \theta). \tag{24}$$

(2) $H(x,\theta)$ is a function for which (11) holds. Then

$$E_R(H(x,\theta)) = E_f(H(x,\theta)). \tag{25}$$

<u>Proof:</u>

We first verify that the hypotheses of Theorem 2.1. are satisfied. We begin with the verification of Assumptions 1-11 of Section 2.1.

<u>Assumption 1.</u>

$\underline{\bar{X}} = R_n$, the n-dimensional Euclidean space.

$\beta_{\underline{\bar{X}}} = B_n$, the n-dimensional Borel field.

For $X \epsilon B_n$,

$$P^\theta(X) = \int_X f(x_1 - \theta, \ldots, x_n - \theta) dx_1 \ldots dx_n. \tag{26}$$

<u>Assumption 2.</u>

$\mathcal{T} = R_1$, the real line.

$\mathcal{A} = R_{n-1}$, the(n-1) dimensional Euclidean space.

In the correspondence, $x \longleftrightarrow (t,a)$,

$$t = x_1 \text{ and } a = (a_1, \ldots, a_{n-1}), \tag{27}$$

where $a_i = x_{i+1} - x_1 \; (i = 1, \ldots, n-1)$.

<u>Assumption 3.</u>

$(\mathcal{T}, \beta_{\mathcal{T}}) = (R_1, B_1)$, $(\mathcal{A}, \beta_{\mathcal{A}}) = (R_{n-1}, B_{n-1})$, $(\Omega, \beta_\Omega) = (R_1, B_1)$

and $\beta_{\mathcal{T}} \times \mathcal{A} = B_n$.

Assumption 4.

g on $\underline{\overline{X}}$ is defined as in (21).

$(\mathcal{G}, \beta_{\mathcal{G}})$ is a measurable space which is in one-to-one correspondence with the measurable space (R_1, B_1). Also, the measure element $d\mu$ is given by dg. Since, Lebesgue measure is invariant under translation, it follows that $\mu(gG) = \mu(G)$ for $g \epsilon \mathcal{G}$ and $G \epsilon \beta_{\mathcal{G}}$.

Assumption 5.

The class of measures P^{θ}, $-\infty < \theta < \infty$, is closed under $\mathcal{G}$. For any $g \epsilon \mathcal{G}$, the variate X with which measure P^{θ} is associated is transformed into a variate gX with which measure $P^{\theta - g}$ is associated, and the later measure is in the class.

Assumption 6.

Also from (23), it follows that there is a unique g^* such that $g^* \theta_1 = \theta_2$, namely $g^* = \theta_1 - \theta_2$.

Assumption 7.

By use of (22), we have that:

$$t_g = x_1 - g \text{ and } a = (x_2 - x_1, \ldots, x_n - x_1). \tag{28}$$

Clearly, t_g depends only on t and g and not on a.

Assumption 8.

The verification of this assumption is similar to that of Assumption 5.

Assumption 9.

This follows from part (1) in the Assumptions of the theorem.

Assumption 10.

It has already been seen that $d\mu(t) = dt$. Take, $d\lambda(a) = da_1 \ldots da_{n-1}$ and $h(t,a) = 1$. Then it is seen that for Sand $S' \epsilon B_n$,

$$\int_S dL_n = \int_{S'} d\mu(t) d\lambda(a), \tag{29}$$

where Sand S' are related through (2(b)).

Assumption 11.

If $T \epsilon \beta_{\mathcal{T}}$ is such that $\int_T d\mu(t) = \int_T dt = 0$, then by use of (29) we have that for any $A \epsilon \beta_{\mathcal{A}}$, $P^{\omega}(T \times A) = 0$. From (14), it follows $P_1^{\theta}(T \mid a) = 0$. Hence, $P_1^{\theta} << \mu$.

Also, $p_1(t \mid a, \theta)$, the conditional density of t with respect to the measure μ is given by:

$$p_1(t \mid a, \omega) = \frac{f(t - \theta, a_1 + t - \theta, \ldots, a_{n-1} + t - \theta)}{\int_{\mathcal{T}} f(t - \theta, a_1 + t - \theta, \ldots, a_{n-1} + t - \theta) dt}. \tag{30}$$

The denominator equals the marginal density of a, and therefore the integral exists.

Finally, we have that:

$\Delta(t) = 1$ and $d\nu(\theta) = d\theta$.

The fiducial density of θ given x by use of (19) is:

$$p(\theta \mid x) = p_1(t \mid a, \omega). \tag{31}$$

Thus we see that all the hypotheses of Theorem 2.1. are satisfied and consequently (25) holds.

For the proofs of the following Corollaries, the reader is referred to Pitman (1939).

Corollary 2.1.

If Assumption (1) of Theorem 2.2. is satisfied then the fiducial mean which we denote by $\hat{\theta}_M$ is the minimum mean square error invariant estimator.

Corollary 2.2.

If Assumption (1) of Theorem 2.2. is satisfied, then the fiducial median which we denote by $\hat{\theta}_C$ is an estimator such that:

$$E(|\hat{\theta}_C - \theta|) \leq E(|\hat{\theta} - \theta|), \tag{32}$$

where $\hat{\theta}$ is any other invariant estimator.

Definition 2.6.

An estimator $\hat{\theta}_B$ is the best estimator of θ (in the Pitman sense) if for all the positive values of h we have that:

$$P\{|\hat{\theta}_B - \theta| \leq h\} \geq P\{|\hat{\theta} - \theta| \leq h\} \tag{33(a)}$$

and for some positive value of h,

$$P\{|\hat{\theta}_B - \theta| \leq h\} > P\{|\hat{\theta} - \theta| \leq h\} \tag{33(b)}$$

Corollary 2.3.

Let the fiducial distribution $p(\theta \mid x)$ be unimodal (with mode at $\hat{\theta}_L$). Also, let $p(\theta_1 \mid x) \leq p(\theta_2 \mid x)$, for $-\infty < \theta_1 \leq \theta_2 \leq \theta_M$ and $p(\theta_1 \mid x) \geq p(\theta_2 \mid x)$ for $\theta_M < \theta_1 \leq \theta_2 < \infty$, then $\hat{\theta}_M = \hat{\theta}_C = \hat{\theta}_L = \hat{\theta}_B$.

2.4. Case of One Location Parameter and One Scale Parameter.

In this section we consider a theorem which deals with families of distributions with one location parameter θ and one scale parameter σ.

Definition 2.8.

An estimator $\hat{\theta}(x_1, \ldots, x_n)$ is said to be an invariant estimator of θ if for all a, $-\infty < a < \infty$ and $b > 0$, we have that:

$$\hat{\theta}(b(x_1-a), \ldots, b(x_n-a)) = b(\hat{\theta}(x_1, \ldots, x_n)-a). \tag{34}$$

Definition 2.9.

An estimator $\hat{\sigma}(x_1, \ldots, x_n)$ is said to be an invariant estimator of σ if for all a, $-\infty < a < \infty$, and b, $b > 0$, we have that:

$$\hat{\sigma}(b(x_1-a), \ldots, b(x_n-a)) = b(\hat{\sigma}(x_1, \ldots, x_n)). \tag{35}$$

Definition 2.10.

For $\underline{X} = R_n$, define $g_{a,b}$ on $\underline{X}$ as follows:

$$g_{a,b}\, x = (b(x_1-a), \ldots, b(x_n-a)). \tag{36}$$

For $\Omega = \{(a,b)\}$, $-\infty < a < \infty$, $0 < b < \infty$, define $g^*_{a,b}$ as follows:

$$g^*_{a,b}\,(\theta, \sigma) = (b(\theta-a), b\sigma). \tag{37}$$

Pitman (1939) has considered this case and has shown the identity (11) to hold for functions of the form $\varphi\left[\frac{\hat{\theta}(x_1, \ldots, x_n)-\theta}{\sigma}\right]$ and $\varphi\left[\frac{\hat{\sigma}(x_1, \ldots, x_n)}{\sigma}\right]$ where $\hat{\theta}(x_1, \ldots, x_n)$ and $\hat{\sigma}(x_1, \ldots, x_n)$ are invariant estimators of θ and σ respectively in the sense of (34) and (35) respectively. A noteworthy feature of the following theorem is that it contains an example of a discrete ancillary, namely a_0. The values taken by a_0 are 1 and 0 according as $x_2 \geq x_1$ or $x_2 < x_1$.

Theorem 2.3.

Let $a_0 = 1$ or 0 according as $x_2 \geq x_1$ or $x_2 < x_1$ and let $a_i = \frac{x_{i+2}-x_1}{x_2-x_1}$ $(i = 1, \ldots, n-2)$. Also, let E_R denote the conditional expectation over a region R in which the ancillaries $a_0, a_1, \ldots, a_{n-2}$ are fixed and E_f denote the expectation with respect to the fiducial density of θ, σ given x.

Assume that:

(1) x has density with respect to Lebesgue measure L_n given by:

$$p(x, \theta, \sigma) = \sigma^{-n} f\left(\frac{x_1-\theta}{\sigma}, \ldots, \frac{x_n-\theta}{\sigma}\right), \ -\infty < x_1, \ldots, x_n, \theta < \infty, \ \sigma > 0. \qquad (38)$$

(2) $H(x, \theta, \sigma)$ is a function for which (11) holds. Then

$$E_R(H(x, \theta, \sigma)) = E_f((H_x, \theta, \sigma)). \qquad (39)$$

Proof:

The proof of this theorem is similar to that of Theorem 2.3. The verification of the assumptions will be gone into briefly.

Assumption 2.

In the correspondence x <--> (t, a), t and a are given by:

$t = (t_1, t_2)$ where $t_1 = x_1$ and $t_2 = |x_2-x_1|$, and

$a = (a_0, a_1, \ldots, a_{n-2})$, where $a_0 = 1$ if $x_2 \geq x_1$ and $a_0 = 0$ if $x_2 < x_1$ and $a_i = \frac{x_{i+2}-x_1}{x_2-x_1}$, $(i = 1, \ldots, n-2)$.

Assumption 4.

The definition of $g_{a;b}$ on $\underline{\overline{X}}$ is given by (35).
The measure element of the measure μ (left Haar measure) is $\frac{da\,db}{b^2}$.

Assumption 10.

The measure elements $d\mu(t)$ and $d\lambda(a)$ are given by:

$$d\mu(t) = \frac{dt_1 dt_2}{t_2^2} \quad \text{and} \quad d\lambda(a) = da_1 \ldots da_{n-2}\,d\eta(a_0), \tag{40}$$

where η is a discrete measure having mass 1/2 at each of the two values $a_0 = 0, 1$. For $S \epsilon \beta_{\underline{\overline{X}}}$ and $S' \epsilon \beta_{T \times A}$, it can be shown that the following holds:

$$\int_{T \times A} 2t_2^n \, d\mu(t) \, d\lambda(a) = \int_S dL_n(x), \tag{41}$$

where S and S' are related through (2(b)). Also, $p_1(t|a,\omega)$, the conditional density of t with respect to the measure μ is given by:

$$p_1(t|a, \omega) = \frac{t_2^n \, f\left(\frac{x_1-\theta}{\sigma}, \ldots, \frac{x_n-\theta}{\sigma}\right)}{\int_T t_2^n \, f\left(\frac{x_1-\theta}{\sigma}, \ldots, \frac{x_n-\theta}{\sigma}\right) d\mu(t)}. \tag{42}$$

Finally we have that:

$$\Delta(t) = \frac{1}{t_2} \quad \text{and} \quad d\nu(\theta, \sigma) = \frac{d\theta d\sigma}{\sigma}$$

Hence, by use of (19) $p(\theta,\sigma|x)$ the fiducial density of θ,σ given x is given by:

$$p\,\theta,\sigma\,|\,x) = \frac{|x_2-x_1|^{n-1} f\left(\frac{x_1-\theta}{\sigma}, \ldots, \frac{x_n-\theta}{\sigma}\right)}{\int_T t_2^n \, f\left(\frac{x_1-\theta}{\sigma}, \ldots, \frac{x_n-\theta}{\sigma}\right) d\mu(t)}. \tag{43}$$

Thus we see that all the hypotheses of Theorem 2.1 are satisfied and so we have that (39) holds.

Lemma 2.5.

Let $H_1(x,\theta,\sigma) = \varphi\left[\frac{\hat{\theta}(x_1, \ldots, x_n)-\theta}{\sigma}\right]$ and $H_2(x,\theta,\sigma) = \varphi\left[\frac{\hat{\sigma}(x_1, \ldots, x_n)}{\sigma}\right]$, where $\hat{\theta}(x_1, \ldots, x_n)$ and $\hat{\sigma}(x_1, \ldots, x_n)$ are functions for which (34) and (35) hold respectively. Then for $H_1(x,\theta,\sigma)$ and $H_2(x,\theta,\sigma)$, (11) holds.

Proof:

$$\begin{aligned}
H_1(g_{a,b}\, x, g^*_{a,b}(\theta,\sigma)) &= \varphi\left[\frac{\hat{\theta}(b(x_1-a), \ldots, b(x_n-a)) - b(\theta-a)}{b\sigma}\right], \\
&= \varphi\left[\frac{b(\hat{\theta}(x_1, \ldots, x_n) - a) - b(\theta-a)}{b\sigma}\right], \text{ by (34)}, \\
&= \varphi\left[\frac{\hat{\theta}(x_1, \ldots, x_n) - \theta}{\sigma}\right], \\
&= H_1(x,\theta,\sigma).
\end{aligned}$$

Thus, (11) holds.

$$\begin{aligned}
H_2(g_{a,b}\, x, g^*_{a,b}(\theta,\sigma)) &= \varphi\left[\frac{\hat{\sigma}(b(x_1-a), \ldots, b(x_n-a))}{b\sigma}\right], \\
&= \varphi\left[\frac{b\hat{\sigma}(x_1, \ldots, x_n)}{b\sigma}\right], \text{ by (35)}, \\
&= \varphi\left[\frac{\hat{\sigma}(x_1, \ldots, x_n)}{\sigma}\right].
\end{aligned}$$

Thus, (11) again holds.

For the proof of corollaries 2.4., 2.5., and 2.6., the reader is referred to Pitman (1939).

Corollary 2.4.

If Assumption (1) of Theorem 2.3 is satisfied, then $E_f(\theta/\sigma^2)/E_f(1/\sigma^2)$ is the minimum mean square error invariant estimator of θ.

Corollary 2.5.

If Assumption (1) of Theorem 2.3 is satisfied, then $\hat{\theta}_c$ (the median of the marginal fiducial distribution of θ) is the closest invariant estimator of θ i.e.

$$P(|\hat{\theta}_c - \theta|) \leq P(|\hat{\theta} - \theta|), \tag{44}$$

where $\hat{\theta}$ is any invariant estimator.

Corollary 2.6.

If Assumption (1) of Theorem 2.3 is satisfied, then $E_f(\frac{1}{\sigma})/E_f(\frac{1}{\sigma^2})$ is the minimum mean square error invariant estimator of σ.

Remark 2.10.

Since $\frac{(\hat{\theta} - \theta)^2}{\sigma^2}$ and $\frac{|\hat{\theta} - \theta|}{\sigma}$ are functions of the form $H_1(x, \theta, \sigma)$ and $H_2(x, \theta, \sigma)$ respectively, we have by Lemma 2.5, that for them (11) holds. Hence, Corollaries 2.4, 2.5, and 2.6 are special cases of Theorem 2.3.

2.5. Case of Two Location Parameters and One Scale Parameter.

In this section we consider a theorem which deals with families of distributions with two location parameters and one scale parameter. Again, Pitman (1939) has considered this case, but has assumed that $x_1, \ldots, x_m, y_1, \ldots, y_n$ are mutually independent. A noteworthy feature of the theorem is that it contains an example of a new ancillary, namely $c = \dfrac{y_2-y_1}{x_2-x_1}$.

Definition 2.11.

For $\underline{\bar{X}} = R_{m+n}$, define $g_{a_1,a_2,b}$ on $\underline{\bar{X}}$ as follows:

$$g_{a_1,a_2,b} \quad x = (b(x_1-a_1), \ldots, b(x_m-a_1), b(y_1-a_2), \ldots, b(y_n-a_2)). \tag{45}$$

For $\Omega = \{(a_1,a_2,b)\}$, $-\infty < a_1 < \infty$, $-\infty < a_2 < \infty$, $0 < b < \infty$, define $g^*_{a_1,a_2,b}$ as follows:

$$g^*_{a_1,a_2,b} (\theta_1,\theta_2, \sigma) = (b(\theta_1-a_1), b(\theta_2-a_2), b\sigma). \tag{46}$$

Theorem 2.4.

Let $a_0 = 1, 0$ according as $x_2 \geq x_1$ or $x_2 < x_1$, $a_i = \dfrac{x_{i+2}-x_1}{x_2-x_1}$ $(i = 1, \ldots, m-2)$, $b_j = \dfrac{y_{j+2}-y_1}{y_2-y_1}$ $(j = 1, \ldots, n-2)$ and $c = \dfrac{y_2-y_1}{x_2-x_1}$.

Also let E_R denote the conditional expectation over a region R in which the ancillaries $a_0, a_1, \ldots, a_{m-2}, b_1, \ldots, b_{n-2}$ and c are fixed and E_f denote the expectation with respect to the fiducial density of θ_1, θ_2, σ given x and y.

Assume:

(1) x, y has density with respect to Lebesgue measure L_{m+n} given by:

$$p\ (x,\ y,\ \theta_1,\ \theta_2,\ \sigma) = \sigma^{-(m+n)}\ f\left(\frac{x_1-\theta_1}{\sigma},\ \ldots,\ \frac{x_m-\theta_1}{\sigma},\ \frac{y_1-\theta_2}{\sigma},\ \ldots,\ \frac{y_n-\theta_2}{\sigma}\right). \qquad (47)$$

(2) $H(x, y, \theta_1, \theta_2, \sigma)$ is a function for which (11) holds. Then

$$E_R\ (H(x_1 y,\ \theta_1,\ \theta_2,\ \sigma)) = E_f\ (H(x,\ y,\ \theta_1,\ \theta_2,\ \sigma)). \qquad (48)$$

Proof:

The proof of this theorem is similar to the proof of Theorems 2.2 and 2.3 and so it will not be given.

Definition 2.12.

An estimator $\hat{\gamma}(x_1, \ldots, x_m, y_1, \ldots, y_n)$ of $\gamma = \theta_1 - \theta_2$ will be called invariant if for all a_1, a_2, b, $-\infty < a_1, a_2 < \infty$, $b > 0$ we have that:

$$\hat{\gamma}(b(x_1-a_1),\ \ldots,\ b(x_m-a_1),\ b(y_1-a_2),\ \ldots,\ b(y_n-a_2))$$
$$= b(\hat{\gamma}(x_1,\ \ldots,\ x_m,\ y_1,\ \ldots,\ y_n)\ -\ (a_2-a_1)). \qquad (49)$$

Lemma 2.6.

Let $H(x, y, \theta_1, \theta_2, \sigma) = \varphi\left[\frac{\hat{\gamma}(x_1,\ \ldots,\ x_m,\ y_1,\ \ldots y_n) - \gamma}{\sigma}\right]$

where $\hat{\gamma}(x_1, \ldots, x_n, y_1, \ldots, y_n)$ is a function for which (49) holds. Then $H(x, y, \theta_1, \theta_2, \sigma)$ satisfies (11).

Proof:

The proof of this is straightforward and will not be given.

Corollary 2.7.

If Assumption (1) of Theorem 2.4 is satisfied, then

$E_f(\frac{\gamma}{\sigma^2}) / E_f(\frac{1}{\sigma^2})$ is the minimum mean square error invariant estimator of γ.

Proof:

By use of Lemma 2.6, we see that (11) holds for $(\frac{\hat{\gamma}-\gamma}{\sigma^2})^2$. Hence, by Theorem 2.4, we have that:

$$E_R \left(\frac{\hat{\gamma}-\gamma}{\sigma}\right)^2 = E_f \left(\frac{\hat{\gamma}-\gamma}{\sigma}\right)^2. \tag{50}$$

Since the right hand side of (50) minimum implies $0 = E_f \left(\frac{\hat{\gamma}-\gamma}{\sigma^2}\right) = \hat{\gamma} E_f \left(\frac{1}{\sigma^2}\right) = E_f \left(\frac{\gamma}{\sigma^2}\right)$. Hence $\hat{\gamma} = E_f \left(\frac{\gamma}{\sigma^2}\right) / E_f \left(\frac{1}{\sigma^2}\right)$. Consequently, $E_f \left(\frac{\gamma}{\sigma^2}\right) / E_f \left(\frac{1}{\sigma^2}\right)$ is the minimum mean square error invariant estimator.

Remark 2.11.

A general identity of the form (48) was not considered by Pitman (1939).

2.6. Case of Two Location Parameters and Two Scale Parameters.

In this section we consider a theorem with families of distributions with two location parameters and two scale parameters. For this case we have not been able to obtain or even define in a satisfactory way an invariant estimator for the difference of two location parameters.

Definition 2.13.

For $\underline{\overline{X}} = R_{m+n}$ define g_{a_1,b_1,a_2,b_2} on $\underline{\overline{X}}$ as follows:

$$g_{a_1,b_1,a_2,b_2}(x,y) = (b_1(x_1-a_1), \ldots, b_1(x_m-a_1), b_2(y_1-a_2), \ldots, b_2(y_n-a_2)). \tag{51}$$

For $\Omega = \{(a_1,b_1,a_2,b_2)\}$, $-\infty < a_1 < \infty$, $0 < b_1 < \infty$, $-\infty < a_2 < \infty$, $0 < b_2 < \infty$, define $g^*_{a_1,b_1,a_2,b_2}$ as follows:

$$g^*_{a_1,b_1,a_2,b_2}(\theta_1, \sigma_1, \theta_2, \sigma_2) = (b_1(\theta_1-a_1), b_1\sigma_1, b_2(\theta_2-a_2), b_2\sigma_2) \tag{52}$$

Theorem 2.5.

Let $a_0 = 1, 0$ according as $x_2 \geq x_1$ or $x_2 < x_1$ and $b_0 = 1, 0$ according as $y_2 \geq y_1$ or $y_2 < y_1$, let $a_i = \dfrac{x_{i+2}-x_1}{x_2-x_1}$ $(i = 1, \ldots, m-2)$ and $b_j = \dfrac{y_{j+2}-y_1}{y_2-y_1}$ $(j = 1, \ldots, n-2)$. Also let E_R denote the conditional expectation over a region R in which the ancillaries $a_0, a_1, \ldots, a_{m-2}$, $b_0, b_1, \ldots, b_{n-2}$ are fixed and E_f denote the expectation with respect to the fiducial density of $\theta_1, \theta_2, \sigma_1, \sigma_2$, given x and y.

Assume that:

(1) x,y has density with respect to the Lebesgue measure L_{m+n} given by:

$$p\,(x,\, y,\, \theta_1,\, \sigma_1,\, \theta_2,\, \sigma_2) = \sigma_1^{-m}\, \sigma_2^{-n}\, f\left(\frac{x_1-\theta_1}{\sigma_1},\, \ldots,\, \frac{x_m-\theta_1}{\sigma_1},\, \frac{y_1-\theta_2}{\sigma_2},\, \ldots,\, \frac{y_n-\theta_2}{\sigma_2}\right) \tag{53}$$

(2) $H(x, y, \theta_1, \sigma_1, \theta_2, \sigma_2)$ is a function for which (11) holds. Then

$$E_R\,(H(x,\, y,\, \theta_1,\, \sigma_1,\, \theta_2,\, \sigma_2)) = E_f(H(x,\, y,\, \theta_1,\, \sigma_1,\, \theta_2,\, \sigma_2)) \tag{54}$$

Proof:

The proof of this theorem is similar to proofs of Theorems 2.2, 2.3, 2.4 and so will not be given.

We now proceed to prove a lemma, which will give two classes of functions for which (54) holds. Finally, we obtain a minimum mean square error invariant estimator for $\gamma = \frac{\sigma_1}{\sigma_2}$.

Lemma 2.7.

Let:

$$H(x,\, y,\, \theta_1,\, \theta_2,\, \sigma_1,\, \sigma_2) = \varphi\left[\frac{\hat{\theta}_1(x_1,\, \ldots,\, x_m) - \theta_1}{\sigma_1},\, \frac{\hat{\theta}_2(y_1, \ldots,\, y_n) - \theta_2}{\sigma_2}\right] \tag{55}$$

and

$$G(x,\, y,\, \theta_1,\, \theta_2,\, \sigma_1,\, \sigma_2) = \varphi\left[\frac{\hat{\sigma}_1(x_1,\, \ldots,\, x_m)}{\sigma_1},\, \frac{\hat{\sigma}_2(y_1,\, \ldots,\, y_n)}{\sigma_2}\right], \tag{56}$$

where $\hat{\theta}_1(x_1, \ldots, x_m)$ and $\hat{\theta}_2(y_1, \ldots, y_n)$ are invariant estimators of θ_1 and θ_2 respectively in the sense of (34) and $\hat{\sigma}_1(x_1, \ldots, x_m)$ and $\hat{\sigma}_2(y_1, \ldots, y_n)$ are invariant estimators of σ_1 and σ_2 respectively in the sense of (35).

Then (11) holds for $H(x, y, \theta_1, \theta_2, \sigma_1, \sigma_2)$ and $G(x, y, \theta_1, \theta_2, \sigma_1, \sigma_2)$.

Proof:

The proof of this lemma is straight forward and will not be given.

Remark 2.12.

If Assumption (1) of Theorem 2.5. is satisfied, then (54) holds for $H(x,y,\theta_1,\theta_2,\sigma_1,\sigma_2,)$ and $G(x,y,\theta_1,\theta_2,\sigma_1,\sigma_2)$.

Definition 2.14.

An estimator $\hat{\rho}(x_1, \ldots, x_m, y_1, \ldots, y_n)$ will be said to be ratio-invariant estimator of $\rho = \frac{\sigma_1}{\sigma_2}$ if:

$$\hat{\rho}(x_1, \ldots, x_m, y_1, \ldots, y_n) = \frac{\hat{\rho}_1(x_1, \ldots, x_m)}{\hat{\rho}_2(y_1, \ldots, y_n)}, \tag{57}$$

where $\hat{\rho}_1(x_1, \ldots, x_m)$ and $\hat{\rho}_2(y_1, \ldots, y_n)$ are invariant functions in the sense of (35).

Corollary 2.8.

If Assumption (1) of Theorem 2.5. is satisfied, then $E_f(\frac{1}{\rho}) \,/\, E_f(\frac{1}{\rho^2})$ is the minimum mean square ratio-invariant estimator of ρ.

Proof:

Since $\frac{\hat{\rho}(x_1, \ldots, c_m, y_1, \ldots, y_n)}{\rho}$ is a function of the form (56), we have by the Remark 2.12, that (54) holds for it. In particular, (54) holds for $(\frac{\hat{\rho}}{\rho} - 1)^2$. We want to determine a $\hat{\rho}(x_1, \ldots, x_m, y_1, \ldots, y_n)$ in the class of ratio-invariant estimators such that $E(\hat{\rho} - \rho)^2$ is minimized. Equivalently, we want to obtain a $\hat{\rho}(x_1, \ldots, x_m, y_1, \ldots, y_n)$ such that

$$E\left[\frac{\hat{\rho}(x_1, \ldots, x_m, y_1, \ldots, y_n)}{\rho} - 1\right]^2$$ is minimized. By an argument similar to that used in Corollary 2.7., we have that $\hat{\rho} = E_f(\frac{1}{\rho}) / E_f(\frac{1}{\rho^2})$.

2.7. Some Remarks.

Several other cases have not been considered, e.g., the cases of (1) two location parameters, and (2) k location parameters and k scale parameters $(k > 2)$. We want to mention here that it is easily seen that these cases are also special cases of Theorem 2.1. Thus, they can also be handled in an analogous manner.

Chapter 3. Estimation in the Case of Rotation Families.

3.0. Introduction.

In this Chapter, we consider the problem of estimation of the angle α through which the axes have been rotated. It will be seen that Assumptions 1-11 of Chapter 2 are satisfied for the specification of the problem. Accordingly, it will be possible to make use of the expectation identity (20) of Chapter 2 to obtain certain "best" estimators. Since the parameter α is an angle, its fiducial distribution (when defined) is defined on a circle. Some non-fiducial theory of circular distributions has been recently considered by Breitenburger (1963) who lists about thirty references to earlier work. A much longer bibliography can be found in Greenwood (1959).

3.1. The Formulation of the Problem.

The formulation of the problem is as follows:

Let $X = (X_1, X_2)$ be a random variable and suppose we have available to us n bivariate observations $x_i = (x_{i1}, x_{i2})$ $(i = 1,2,\ldots,n)$. Let $x = (x_1, \ldots, x_n)$. Assume that the density of x is $f(x)$ and it has no parameters.

Consider the following transformations $\mathcal{G} = \{g_\alpha : \ 0 \leq \alpha < 2\pi\}$, where

$$g_\alpha(x) = (x_{11} \operatorname{Cos} \alpha + x_{12} \operatorname{Sin} \alpha, -x_{11} \operatorname{Sin} \alpha + x_{12} \operatorname{Cos} \alpha, \ldots, -x_{n1} \operatorname{Cos} \alpha + x_{n2} \operatorname{Sin} \alpha, -x_{n1} \operatorname{Sin} \alpha + x_{n2} \operatorname{Cos} \alpha). \quad (1)$$

Then $\{f(x,\alpha) : 0 \leq \alpha < 2\pi\}$ is a family of distributions. They correspond to the family of random variables which would be generated if the axes were rotated and it is clear that the family can be parametrized by $\{\alpha : 0 \leq \alpha < 2\pi\}$. Now consider the same specification in terms of polar coordinates i.e., we make the following transformation:

$$R_i = \sqrt{(X_{i1}^2 + X_{i2}^2)}, \quad \Theta_i = \tan^{-1} \frac{X_{i2}}{X_{i1}}, \quad (2)$$

(where Θ_i, lies in the range $0 \leq \Theta_i < 2\pi$ and is taken to be in the appropriate quadrant depending on the signs of X_{i2} and X_{i1})

$i = 1, \ldots, n$.

Then $\{f_{R,\Theta}(r,\theta,\alpha) : 0 \leq \alpha < 2\pi\}$ is a family of distributions, where $R = (R_1, \ldots, R_n)$, $\Theta = (\Theta_1, \ldots, \Theta_n)$, $r = (r_1, \ldots, r_n)$ and $\theta = (\theta_1, \ldots, \theta_n)$ and $f_{R,\Theta}(r,\theta,\alpha)$ is given by:

$$f_{R,\Theta}(r,\theta,\alpha) = \prod_{i=1}^{n} r_i f(r_1 \operatorname{Cos}(\theta_1 - \alpha), r_1 \operatorname{Sin}(\theta_1 - \alpha), \ldots, r_n \operatorname{Cos}(\theta_n - \alpha), r_n \operatorname{Sin}(\theta_n - \alpha)). \quad (3)$$

In the polar coordinate system the transformations $\mathcal{G}$ can be expressed as follows:

$$g_\alpha(r,\theta) = (r_1, \ldots, r_n, \theta_1 - \alpha, \ldots, \theta_n - \alpha), \tag{4}$$

where $r = (r_1, \ldots, r_n)$ and $\theta = (\theta_1, \ldots, \theta_n)$.

The corresponding group of transformations induced on the parameter space is $\mathcal{G}^* = \{g_\alpha^* : 0 \leq \alpha < 2\pi\}$, where g_α^* is defined as follows:

$$g_\alpha^* (\varphi) = \alpha + \varphi \text{ (Mod. } 2\pi). \tag{5}$$

Our problem is to obtain certain "best" estimators of α, the only parameter which appears in $f_{R,\Theta}(r,\theta,\alpha)$.

<u>Definition 3.1.</u>

Any non-negative function $f(\beta)$ such that $\int_{\beta_0}^{\beta_0+2\pi} f(\beta)d\beta = 1$,

$(-\infty < \beta_0 < \infty)$ will be called a density on a circle.

<u>Remark 3.1.</u>

$f(\beta)$ is an a.e. periodic function of β with period 2π, i.e., $f(\beta) = f(\beta + 2\pi)$ a.e. Lebesgue, $(-\infty < \beta < \infty)$.

Obviously here we are concerned with distributions which are defined over a circle. Such distributions pose special problems for many of the standard statistical tools fail to be meaningful. For example, the usual linearity property of expectation, i.e., $E(a + \beta) = a + E(\beta)$, where $-\infty < a < \infty$, does not necessarily hold. Similarly, the other moments of distribution on the circle are not well defined. Finally, to point out yet another problem, how should one define the difference between two angles. For example, if we refer to Figure 1, should the difference be given by angle 1 or angle 2.

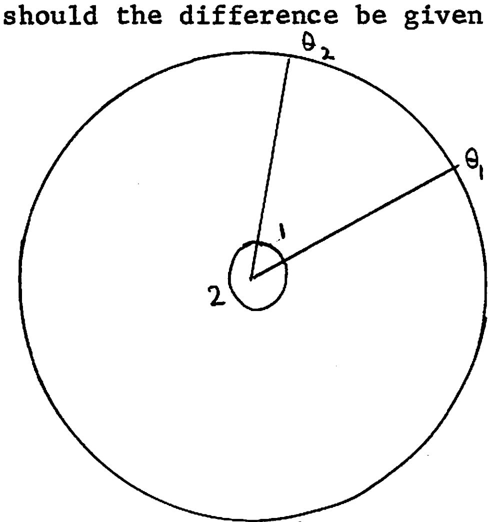

Figure 1.

Thus one has to make meaningful conventions to define such concepts as expectation, difference between two angles, etc.

A reference to a circular fiducial distribution is found in Fisher (1956). He derives the fiducial distribution of α, when X_1 is $N(1,1)$ and X_2 is $N(0,1)$ and X_1 and X_2 are independently distributed. Given one bivariate observation, the fiducial density is found by Fisher to be:

$$f_f(\alpha \mid r,\theta) = \frac{1}{2\pi I_0(r)} e^{r \operatorname{Cos}(\theta-\alpha)}, \tag{6}$$

where I_0 (r) is a modified Bessel function of the first kind.

3.2. Verification of Assumptions.

In this Section we will give a brief account of the verification of the assumptions of Chapter 2 and also give the corresponding definitions for some of the terms involved in those assumptions.

We have the following:

(1) $(\theta_1, \ldots, \theta_n, r_1, \ldots, r_n) \longleftrightarrow (\theta_1, \theta_2 - \theta_1, \ldots, \theta_n - \theta_1,$

$r_1, \ldots, r_n) = (t,a),$

where $t = \theta_1$, $a = (a_1, \ldots, a_{n-1}, r_1, \ldots, r_n)$ and

$a_i = \theta_{i+1} - \theta_1$, $i = 1, \ldots, n-1$.

(2) It is easily verified that $\mathcal{G}^*$ is exactly transitive on Ω.

(3) The transformations $\mathcal{G}' = \{g'_\alpha : 0 \leq \alpha < 2\pi\}$ are defined as follows: $g'_\alpha(\theta_1) = \theta_1 - \alpha$.

(4) The transitivity of $\mathcal{G}'$ on $\mathcal{T}$ is also easily verified.

(5) It can be easily verified that for $A \in \beta_{[0,2\pi]}$ (where $\beta_{[0,2\pi]}$ is the σ-field of subsets of $[0,2\pi]$, the range space of Θ_1) we have that:

$$P_1(A \mid a_1\alpha) = \frac{\int_A h(\theta_1, a, \alpha)d\theta_1}{\int_0^{2\pi} h(\theta_1, a, \alpha)d\theta_1}, \quad \text{where} \tag{7}$$

$$h(\theta_1, a, \alpha) = f(r_1 \mathrm{Cos}(\theta_1 - \alpha), r_1 \mathrm{Sin}(\theta_1 - \alpha), r_2 \mathrm{Cos}(a_1 + \theta_1 - \alpha), r_2 \mathrm{Sin}(a_1 + \theta_1 - \alpha), \ldots, r_n \mathrm{Sin}(a_{n-1} + \theta_1 - \alpha)).$$

3.3. The Conditional and Fiducial Densities.

We have from (7) that $p_1(\theta_1 \mid a,\alpha)$, the conditional density of θ, given a is as follows:

$$p_1(\theta_1 \mid a,\alpha) = \frac{h(\theta_1, a, \alpha)}{\int_0^{2\pi} h(\theta_1, a, \alpha)d\theta_1}. \tag{8}$$

We also have that $\Delta(\theta_1) = 1$ and $d\nu(\alpha) = d\alpha$. Hence, the fiducial density of α given r, θ by use of (19) of Chapter 2 is given by:

$$p(\alpha \mid \theta_1, a) = p_1(\theta_1 \mid a, \alpha). \tag{9}$$

Example 3.1.

The Conditional and Fiducial Densities in Case of One Bivariate Observation from Exponential Distribution.

Let $X = (X_1,X_2)$, where X_1 and X_2 are two random variables distributed independently of each other with densities e^{-x_1} and e^{-x_2} respectively. Suppose we have available to us one bivariate observation. Then, by (8) and (9), the conditional density of θ_1 given r_1 and the fiducial density of α given r_1, θ_1 are given by:

$$p_1(\theta_1 \mid r_1,\alpha) = \frac{e^{-r_1[\mathrm{Cos}(\theta_1-\alpha) + \mathrm{Sin}(\theta_1-\alpha)]}}{\int_0^{2\pi} e^{-r_1[\mathrm{Cos}(\theta_1-\alpha) + \mathrm{Sin}(\theta_1-\alpha)]} d\theta_1}, \qquad (10)$$

$$r \geq 0 \text{ and } \alpha \leq \theta_1 \leq \alpha + \frac{\pi}{2},$$

$$= 0 \quad , \qquad \text{otherwise.}$$

$$p(\alpha \mid r_1,\theta_1) = \frac{e^{-r_1[\mathrm{Cos}(\theta_1-\alpha) + \mathrm{Sin}(\theta_1-\alpha)]}}{\int_0^{2\pi} e^{-r_1[\mathrm{Cos}(\theta_1-\alpha) + \mathrm{Sin}(\theta_1-\alpha)]} d\theta_1}, \qquad (11)$$

$$r > 0 \text{ and } \theta_1 - \frac{\pi}{2} \leq \alpha \leq \theta_1$$

$$= 0 \quad , \qquad \text{otherwise.}$$

Below are given the figures of the fiducial density of α in the two possible cases that could possibly arise.

Case 1. $\frac{\pi}{2} \leq \theta_1 < 2\pi$.

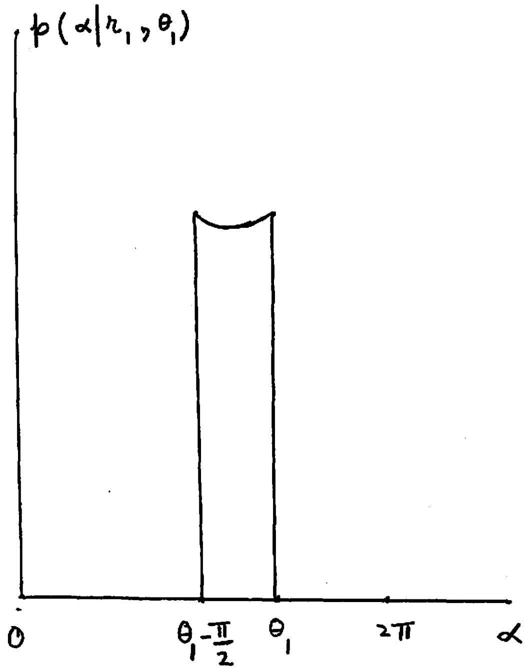

Figure 2(a).

Case 2. $0 \leq \theta_1 < \frac{\pi}{2}$.

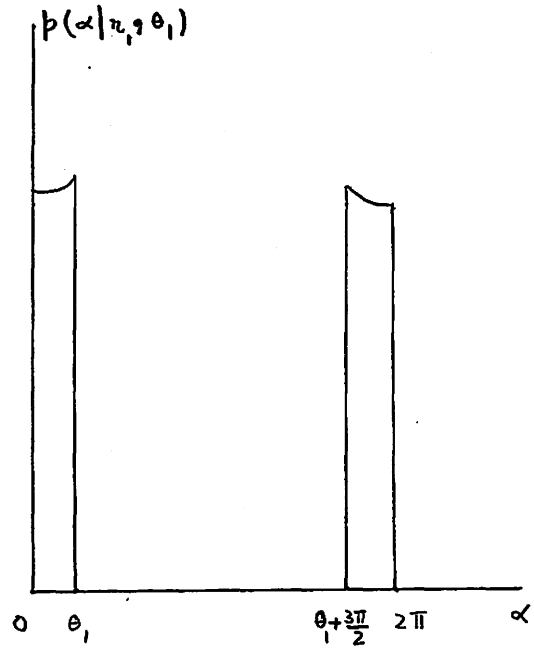

Figure 2(b).

Example 3.2.

The Conditional and Fiducial Densities in Case of Two Bivariate Observations from Exponential Distribution.

Let $X = (X_1, X_2)$ where X_1 and X_2 are two random variables distributed independent of each other with densities e^{-x_1} and e^{-x_2} respectively. Then by (7) and (8), the conditional density of θ_1 given r_1, r_2, a_1 and the fiducial density of α given r_1, r_2, θ_1 and θ_2 are given by:

$$p_1(\theta_1 \mid r_1, r_2, a_1, \alpha) =$$

$$\frac{e^{-r_1[\mathrm{Cos}(\theta_1 - \alpha) + \mathrm{Sin}(\theta_1 - \alpha)]} \cdot e^{-r_2[\mathrm{Cos}(a_1 + \theta_1 - \alpha) + \mathrm{Sin}(a_1 + \theta_1 - \alpha)]}}{\int_0^{2\pi} e^{-r_1[\mathrm{Cos}(\theta_1 - \alpha) + \mathrm{Sin}(\theta_1 - \alpha)]} \cdot e^{-r_2[\mathrm{Cos}(a_1 + \theta_1 - \alpha) + \mathrm{Sin}(a_1 + \theta_1 - \alpha)]} d\theta_1}, \quad (12)$$

$r_1, r_2 > 0$ and the range of $\theta_1, \theta_2 - \theta_1$ is shown in Figure 3.

$= 0$, otherwise.

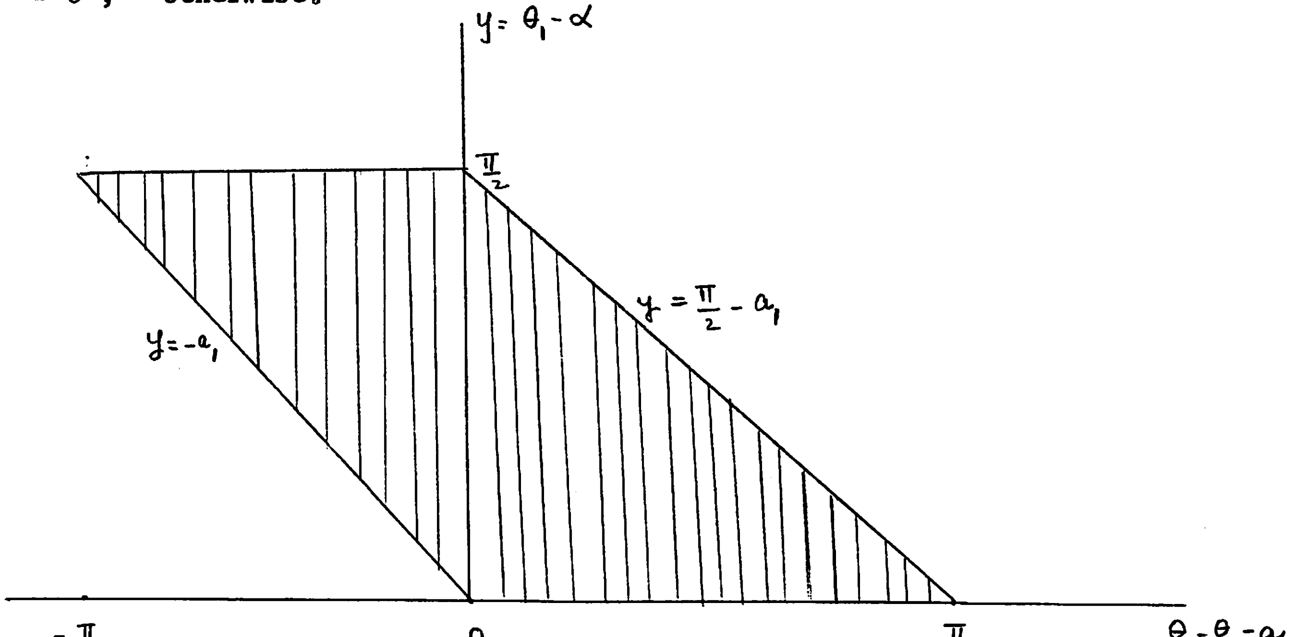

Figure 3.

$$p(\alpha \mid r_1, r_2, \theta_1, \theta_2) = p_1(\theta_1 \mid r_1, r_2, a_1, \alpha), \quad (13)$$

$r_1, r_2 > 0$ and the range of α is determined through Figure 3.

3.4. Invariant Estimators.

3.4.1. Invariant Estimators of α in the Case of One Bivariate Observation.

Definition 3.2.

We will say that $\hat{\alpha}(r_1,\theta_1)$ is an invariant estimator of α if for every φ, $(0 \leq \varphi < 2\pi)$, we have that:

$$\hat{\alpha}(g_\varphi(r_1,\theta_1)) = \hat{\alpha}(r_1,\theta_1) + \varphi, \text{ for } g_\varphi \in \mathcal{G}. \tag{14}$$

The following theorem characterizes invariant estimators of α in case of single bivariate observation.

Theorem 3.1.

An estimator $\hat{\alpha}(r_1,\theta_1)$ is invariant if and only if $\hat{\alpha}(r_1,\theta_1)$ is of the form $(K(r_1) + \theta_1)$, where $K(r_1)$ is some function of r_1.

Proof:

If $\hat{\alpha}(r_1,\theta_1)$ is of the form $K(r_1) + \theta_1$, then obviously it is invariant. Conversely, let $\hat{\alpha}(r_1,\theta_1)$ be any invariant estimator. Differentiating $\hat{\alpha}(r_1,\theta_1)$ with respect to θ_1, we have that:

$$\frac{\partial\hat{\alpha}(r_1,\theta_1)}{\partial\theta_1} = \lim_{h\to 0} \frac{\hat{\alpha}(r_1,\theta_1+h) - \hat{\alpha}(r_1,\theta_1)}{h} = \lim_{h\to 0}\frac{h}{h} = 1, \text{ (due to (14))}.$$

But this implies that any invariant estimator has to be of the form $K(r_1) + \theta_1$.

3.4.2. Invariant Estimators of α in the Case of Two Bivariate Observations.

Following on lines similar to the characterization of invariant estimators of α in case of one bivariate observation, we obtain in this Section a characterization of invariant estimators of α in case of two bivariate observations. Any estimator $\hat{\alpha}$ will be a function of r_1, r_2, θ_1 and θ_2 but it can also be considered to be a function of $r_1, r_2, \theta_2 - \theta_1$ and θ_1.

Definition 3.3.

An estimator $\hat{\alpha}(\theta_1, \theta_2 - \theta_1, r_1, r_2)$ will be called an invariant estimator, if for any φ, $(0 \leq \varphi < 2\pi)$, we have that:

$$\hat{\alpha}(g'_\varphi \theta_1, \theta_2 - \theta_1, r_1, r_2) = \hat{\alpha}(\theta_1, \theta_2 - \theta_1, r_1, r_2) + \varphi, \text{ for } g'_\varphi \in \mathcal{G}'. \quad (15)$$

Theorem 3.2.

An estimator $\hat{\alpha}(\theta_1, \theta_2 - \theta_1, r_1, r_2)$ is invariant if and only if $\hat{\alpha}(\theta_1, \theta_2 - \theta_1, r_1, r_2)$ is of the form $K(r_1, r_2, \theta_2 - \theta_1) + \theta_1$, where $K(r_1, r_2, \theta_2 - \theta_1)$ is some function of $r_1, r_2, \theta_2 - \theta_1$.

Proof:

Differentiating $\hat{\alpha}(\theta_1, \theta_2 - \theta_1, r_1, r_2)$ with respect to the first argument θ_1, we have that

$$\frac{\partial \hat{\alpha}(\theta_1, \theta_2 - \theta_1, r_1, r_2)}{\partial \theta_1} = \lim_{h \to 0} \frac{\hat{\alpha}(\theta_1 + h, \theta_2 - \theta_1, r_1, r_2) - \hat{\alpha}(\theta_1, \theta_2 - \theta_1, r_1, r_2)}{h} =$$

$$\lim_{h \to 0} \frac{h}{h} = 1, \quad \text{(due to (15))}.$$

This implies that $\hat{\alpha}(\theta_1, \theta_2 - \theta_1, r_1, r_2)$ has to be of the form $K(\theta_2 - \theta_1, r_1, r_2) + \theta_1$, where $K(\theta_2 - \theta_1, r_1, r_2)$ is some function of $\theta_2 - \theta_1$, r_1 and r_2. If $\hat{\alpha}(\theta_1, \theta_2 - \theta_1, r_1, r_2)$ is of the form $K(\theta_2 - \theta_1, r_1, r_2) + \theta_1$, then obviously it is an invariant estimator.

Remark 3.2.

Similarly it can be shown that an estimator of α in the case of $n(n > 2)$ bivariate observations is invariant if and only if it is of the form $K(\theta_2 - \theta_1, \ldots, \theta_n - \theta_1, r_1, \ldots, r_n) + \theta_1$, where K is some function of $\theta_2 - \theta_1, \ldots, \theta_n - \theta_1, r_1, \ldots, r_n$.

3.5. Some Definitions.

In Section 3.1, we pointed out that the usual properties of expectation do not necessarily hold when we take expectation with respect to the distributions on the circle. In this Section, we will define certain expectations which will be found to be of interest for obtaining "best" estimators of α. We will also define the concept of symmetry for distributions on the circle.

We begin with the definition of $E_R(\hat{\alpha} - \alpha)^2$, where $\hat{\alpha}$ is any invariant estimator and E_R denotes the expectation with respect to the density (8). Before actually defining this expression, we must make a convention as to what the term $(\hat{\alpha} - \alpha)^2$ means.
We take $(\hat{\alpha} - \alpha)^2$ to be as follows:

$$(\hat{\alpha} - \alpha)^2 = \min_{n=0,1,2,\ldots} (\hat{\alpha} - \alpha \pm 2n\pi)^2 \tag{16}$$

Definition 3.4.

We define $E_R(\hat{\alpha} - \alpha)^2$ as follows:

$$E_R(\hat{\alpha} - \alpha)^2 = \int_{\alpha - \pi - f(a)}^{\alpha + \pi - f(a)} (\hat{\alpha} - \alpha)^2 \, p(\theta_1 \mid a, \alpha) d\theta_1, \tag{17}$$

where $\hat{\alpha} = f(a) + \theta_1$ and $f(a)$ is some function of a.

Remark 3.3.

The reason for this choice of limits is as follows: If $\alpha - \pi - f(a) \leq \theta_1 < \alpha + \pi - f(a)$, then $\alpha - \pi \leq \hat{\alpha} < \alpha + \pi$. Hence, $|\hat{\alpha} - \alpha|$ is always less than or equal to π and consequently $(\hat{\alpha} - \alpha)^2$ is actually equal to the $\min_{n=0,1,2,\ldots} (\hat{\alpha} - \alpha \pm 2n\pi)^2$.

If, $\alpha - \pi - f(a) \leqq \theta_1 < \alpha + \pi - f(a)$, then $\hat{\alpha} - \pi \leqq \alpha < \hat{\alpha} + \pi$. For any invariant estimator $\hat{\alpha}$, we define $E_f(\hat{\alpha} - \alpha)^2$ as follows:

Definition 3.5.

$$E_f(\hat{\alpha} - \alpha)^2 = \int_{\hat{\alpha} - \pi}^{\hat{\alpha} + \pi} (\hat{\alpha} - \alpha)^2 \, p(\alpha \mid r, \theta) d\alpha, \tag{18}$$

where $p(\alpha \mid r,\theta)$ is given by (9).

It may be observed that again in this definition $\mid \hat{\alpha} - \alpha \mid \leqq \pi$ and hence $(\hat{\alpha} - \alpha)^2$ is actually equal to $\min\limits_{n=0,1,2,\ldots} (\hat{\alpha} - \alpha \pm 2n\pi)^2$.

Similar definitions are made for $E_R(\mid \hat{\alpha} - \alpha \mid)$ and $E_f(\mid \hat{\alpha} - \alpha \mid)$.

In a later part of this Chapter, we will need the concept of symmetry and unimodal for distributions on the circle. Consequently, we define these concepts below.

Definition 3.6.

A distribution on the circle is said to be symmetric if there exists an a, $0 \leqq a < 2\pi$ such that $f(a - \gamma) = f(a + \gamma)$, $0 \leqq \gamma < 2\pi$.

Example 3.3.

The distribution referred to in (6) is symmetric around θ. The distribution referred to in (11) is symmetric around $\theta_1 - \frac{\pi}{4}$.

Definition 3.7.

A distribution on the circle will be said to be weakly unimodal if there exists an b_0, $0 \leqq b_0 < 2\pi$ such that $f(b_0) > f(b)$, for all $b \neq 2n\pi + b_0$.

Example 3.4.

Again, the distribution referred to in (6) is unimodal and the mode is at θ.

Theorem 3.3.

If the distribution on the circle is symmetric and unimodal, than it is symmetric around the mode.

Proof:

The proof of this theorem is obvious and will not be given.

3.6. Probability Theory for Distribution on a Circle.

This Section is in three parts. Section 3.6.1. deals with nonfiducial probability theory for distributions on circle. It gives definitions of certain "best" estimators and certain lemmas which follow from these definitions. Section 3.6.2. deals with fiducial probability theory and illustrates the manner in which fiducial argument can be applied to obtain certain "best" invariant estimators. Finally, in Section 3.6.3., the equivalence of fiducial and posterior densities of α is shown when the prior distribution for α is taken to uniform over the interval $(0, 2\pi)$.

3.6.1. Non-Fiducial Probability Theory.

Let $f(\alpha)$ be a circular density in the sense of Section 3.1. Define, $g_0(\hat{\alpha})$ as follows:

$$g_0(\hat{\alpha}) = \int_{\hat{\alpha} - \pi}^{\hat{\alpha} + \pi} (\hat{\alpha} - \alpha)^2 f(\alpha)d\alpha, \quad -\infty < \hat{\alpha} < \infty . \tag{19}$$

Then $g_0(\hat{\alpha})$ is clearly periodic with period 2π. If $g_0(\hat{\alpha})$ is a continuous function of $\hat{\alpha}$, then its minimum exists for it is defined over a circle which is a compact set.

Definition 3.8.

Define $\hat{\alpha}_0$ as follows:

$$\min_{\hat{\alpha}}.\ g_0(\hat{\alpha}) = g_0(\hat{\alpha}_0) , \tag{20}$$

when $\hat{\alpha}_0$ is unique (modulo 2π) it will be called the quasi-mean of the distribution of α.

Also, let $g_1(\hat{\alpha})$ be defined as follows:

$$g_1(\hat{\alpha}) = \int_{\hat{\alpha} - \pi}^{\hat{\alpha} + \pi} | \hat{\alpha} - \alpha | \; f(\alpha) \; d\alpha, \; -\infty < \hat{\alpha} < \infty. \tag{21}$$

Clearly, $g_1(\hat{\alpha})$ is also periodic and if it is a continuous function of $\hat{\alpha}$, its minimum exists.

Definition 3.9.

Define, $\hat{\alpha}_1$ as follows:

$$\min_{\hat{\alpha}} \; g_1(\hat{\alpha}) = g_1(\hat{\alpha}_1), \tag{22}$$

when $\hat{\alpha}$, is unique it will be called the quasi-median of the distribution of α.

Lemma 3.1.

If $\hat{\alpha}_0$ is a quasi-mean and $g_0(\hat{\alpha})$ is differentiable at $\hat{\alpha}_0$, then

$$\int_{\hat{\alpha}_0 - \pi}^{\hat{\alpha}_0 + \pi} (\hat{\alpha}_0 - \alpha) \; f(\alpha) d\alpha = 0. \tag{23}$$

Proof:

Since, by hypothesis the derivative of $g_0(\hat{\alpha})$ exists at $\hat{\alpha}_0$, we have by periodicity of f and by the use of Leibnitz formula that:

$$g_0'(\hat{\alpha}_0) = 2 \int_{\alpha_0 - \pi}^{\hat{\alpha}_0 + \pi} (\hat{\alpha}_0 - \alpha) f(\alpha) d\alpha, \tag{24}$$

since $\hat{\alpha}_0$ is a quasi-mean, it follows that (23) holds.

Lemma 3.2.

If $\hat{\alpha}_1$ is a quasi-median and $g_1(\hat{\alpha}_1)$ is differentiable at $\hat{\alpha}_1$, then

$$\int_{\hat{\alpha}_1-\pi}^{\hat{\alpha}_1} f(\alpha)d\alpha = \frac{1}{2} \tag{25}$$

Proof:

Since by assumption the derivative of $g_1(\hat{\alpha})$ exists at $\hat{\alpha}_1^-$, by use of Leibnitz formula and periodicity of $f(\alpha)$ we have that:

$$g_1'(\hat{\alpha}_1) = \int_{\hat{\alpha}_1-\pi}^{\hat{\alpha}_1} f(\alpha)d\alpha - \int_{\hat{\alpha}_1}^{\hat{\alpha}_1+\pi} f(\alpha)d\alpha . \tag{26}$$

Since, $g_1'(\hat{\alpha}_1) = 0$ and $\int_{\hat{\alpha}_1-\pi}^{\hat{\alpha}_1+\pi} f(\alpha)d\alpha = 1$, it follows that (25) holds.

3.6.2. Fiducial Probability Theory.

Definition 3.8.

An invariant estimator $\hat{\alpha}$ which minimizes $E_f(\hat{\alpha} - \alpha)^2$; where $E_f(\hat{\alpha} - \alpha)^2$ is as defined in (18), will be called $\hat{\alpha}_{MSE}$. Assume that $\hat{\alpha}_{MSE}$ exixts. Since $(\hat{\alpha} - \alpha)^2$ is an invariant function, we have by Theorem 2.1., of Chapter 2, that $E_f(\hat{\alpha} - \alpha)^2 = E_R(\hat{\alpha} - \alpha)^2$ for any invariant estimator $\hat{\alpha}$. Hence, $E_R(\hat{\alpha}_{MSE} - \alpha)^2 \leq E_R(\hat{\alpha} - \alpha)^2$. Consequently, $E(\hat{\alpha}_{MSE} - \alpha)^2 \leq E(\hat{\alpha} - \alpha)^2$, where $\hat{\alpha}$ is any invariant estimator. Thus $\hat{\alpha}_{MSE}$ is a minimum mean square error invariant estimator.

Definition 3.11.

An invariant estimator which minimizes $E_f(| \hat{\alpha} - \alpha |)$, will be called $\hat{\alpha}_C$. Assume that $\hat{\alpha}_C$ exists. In a manner similar to above it can be shown that:

$$E(| \hat{\alpha}_C - \alpha |) \leq E(| \hat{\alpha} - \alpha |). \tag{27}$$

3.6.3. Equivalence of Fiducial and Posterior Distributions of α.

The joint density $f(\theta_1,a,\alpha)$ of (θ_1,a) by using (3), is given by:

$$f(\theta_1,a,\alpha) = \prod_{i=1}^{n} r_i f(r_1 \mathrm{Cos}(\theta_1 - \alpha), r_1 \mathrm{Sin}(\theta_1 - \alpha), r_2 \mathrm{Cos}(a_1 + \theta_1 - \alpha),$$

$$r_2 \mathrm{Sin}(a_1 + \theta_1 - \alpha), \ldots, r_n \mathrm{Sin}(a_{n-1} + \theta_1 - \alpha)). \tag{28}$$

Assume that the prior density of α is uniform over the interval $(0,2\pi)$. Then, the posterior density $f_p(\alpha \mid \theta_1,a)$ of α given θ_1,a is given by:

$$f_p(\alpha \mid \theta_1,a) = \frac{f(\theta_1,a,\alpha)}{\int_0^{2\pi} f(\theta_1,a,\alpha)d\alpha}. \tag{29}$$

Since it can be easily verified that:

$$\int_0^{2\pi} f(\theta_1,a,\alpha)d\theta_1 = \int_0^{2\pi} f(\theta_1,a,\alpha)d\alpha, \tag{30}$$

we have that $f_p(\alpha \mid \theta_1,a) = p_1(\theta_1 \mid a,\alpha)$, which by (9) is the fiducial density of α. This establishes the equivalence of fiducial and posterior densities of α.

3.7. Derivation of "Best" Estimators in Special Cases.

In this Section we will give examples of "best" estimators by applying the theory developed in Sections 3.6.1 and 3.6.2.

Example 3.5.

Case of Exponential Distribution with One Bivariate Observation.

Let $X = (X_1, X_2)$, where X_1 and X_2 are random variables distributed independently of each other with densities e^{-x_1} and e^{-x_2} respectively. Suppose we have available to us one bivariate observation. Also suppose that $r_1 = 1$ and $\theta_1 = \frac{\pi}{2}$. Then by use of (11) we have that the fiducial density of α given $r_1 = 1$ and $\theta_1 = \frac{\pi}{2}$ is given by:

$$p(\alpha \mid r_1 = 1, \theta_1 = \tfrac{\pi}{2}) = \frac{e^{-[\operatorname{Cos}\alpha + \operatorname{Sin}\alpha]}}{\int_0^{\frac{\pi}{2}} e^{-[\operatorname{Cos}\alpha + \operatorname{Sin}\alpha]} d\alpha} \quad \text{for } 0 \leqq \alpha \leqq \tfrac{\pi}{2} \tag{31}$$

$$= 0 \text{ otherwise.}$$

We see that $\hat{\alpha} = \frac{\pi}{4}$ satisfies

$$\int_{\hat{\alpha}-\pi}^{\hat{\alpha}+\pi} (\hat{\alpha} - \alpha)\, p(\alpha \mid r_1 = 1, \theta_1 = \tfrac{\pi}{2}) d\alpha = 0.$$

It is also easily seen that $\hat{\alpha}_0 = \frac{\pi}{4}$ is the minimizing value of the integral

$$\int_{\hat{\alpha}+\pi}^{\hat{\alpha}-\pi} (\hat{\alpha} - \alpha)^2\, p(\alpha \mid r_1 = 1, \theta_1 = \tfrac{\pi}{2}) d\alpha.$$

By use of theory in Section 3.6.2., we have that $\hat{\alpha}_{MSE} = \frac{\pi}{4}$. Thus the minimum mean square error invariant estimator of α takes the value $\frac{\pi}{4}$. It can similarly be shown that $\hat{\alpha}_C = \frac{\pi}{4}$ is an invariant estimator with the property (27).

Example 3.6.

Case of Uniform Distribution with One Bivariate Observation.

Let, $X = (X_1, X_2)$, where X_1 and X_2 are random variables distributed independently of each other and the density of each is uniform over the interval $(0,1)$. Suppose, we have available to us one bivariate observation. Also suppose that $r_1 = \frac{1}{2}$ and $\theta_1 = \frac{\pi}{4}$. Then by use of (11) we have that the fiducial density of α given $r_1 = 1$ and $\theta_1 = \frac{\pi}{2}$ is given by:

$$p(\alpha \mid r_1 = 1, \theta_1 = \tfrac{\pi}{2}) = \tfrac{2}{\pi}, \text{ if } 0 \leq \alpha \leq \tfrac{\pi}{2}, \tag{32}$$

$$= 0 \text{ , otherwise.}$$

We see that $\hat{\alpha} = \frac{\pi}{4}$ satisfies $\int_{\hat{\alpha}-\pi}^{\hat{\alpha}+\pi} (\hat{\alpha} - \alpha)\frac{2}{\pi} d\alpha = 0$. Also, proceeding on lines similar to the above example we have that the minimum mean square invariant estimator of α is $\frac{\pi}{4}$. It also follows that the invariant estimator $\hat{\alpha}_C$ with the property (27) takes the value $\frac{\pi}{4}$.

Example 3.7.

Case of Exponential Distribution with Two Bivariate Observations.

Consider again the problem referred to in Example 2 of Section 3.3.2. Suppose that $r_1 = r_2 = 1$ and $\theta_1 = \theta_2 = \frac{\pi}{2}$ is given by:

$$p(\alpha \mid r_1 = r_2 = 1, \theta_1 = \theta_2 = \tfrac{\pi}{2}) = \frac{e^{-2[\text{Cos } \alpha + \text{Sin } \alpha]}}{\int_0^{\frac{\pi}{2}} e^{-2[\text{Cos } \alpha + \text{Sin } \alpha]} d\alpha}, \tag{33}$$

for $0 \leq \alpha \leq \frac{\pi}{2}$, and equals 0 otherwise.

We see that $\alpha = \frac{\pi}{4}$ satisfies $\int_{\hat{\alpha} - \pi}^{\hat{\alpha} + \pi} (\hat{\alpha} - \alpha) p(\alpha \mid r_1 = r_2 = 1, \theta_1 = \theta_2 = \frac{\pi}{2}) d\alpha = 0.$

Proceeding on lines similar to Example 3.5., we have that the minimum mean square invariant estimator of α is $\frac{\pi}{4}$. It also follows that the invariant estimator α_C with the property (27) takes the value $\frac{\pi}{4}$.

3.8. A Theorem for Symmetric and Unimodal Circular Distributions.

This Section contains a theorem regarding symmetric and unimodal distributions. It deals with the minimization of a certain integral. In the next Section, this theorem will be used to obtain a Bayes' estimate of α.

We begin with proving certain lemmas which will be required in the proof of the theorem.

Lemma 3.3.

Assume:

(1) $f(t)$ and $L(t)$ are non-negative functions defined on $\lambda_1 \leqq t \leqq \lambda_2$.

(2) $f(t)$ is non-increasing on (λ_1, λ_2).

(3) $L(t)$ is non-decreasing on (λ_1, λ_2).

Then,
$$\int_{\lambda_1}^{\lambda_2} f(t)L(t)dt \leqq \int_{\lambda_1}^{\lambda_2} f(t)L(\lambda_1 + \lambda_2 - t)dt. \tag{34}$$

Proof:

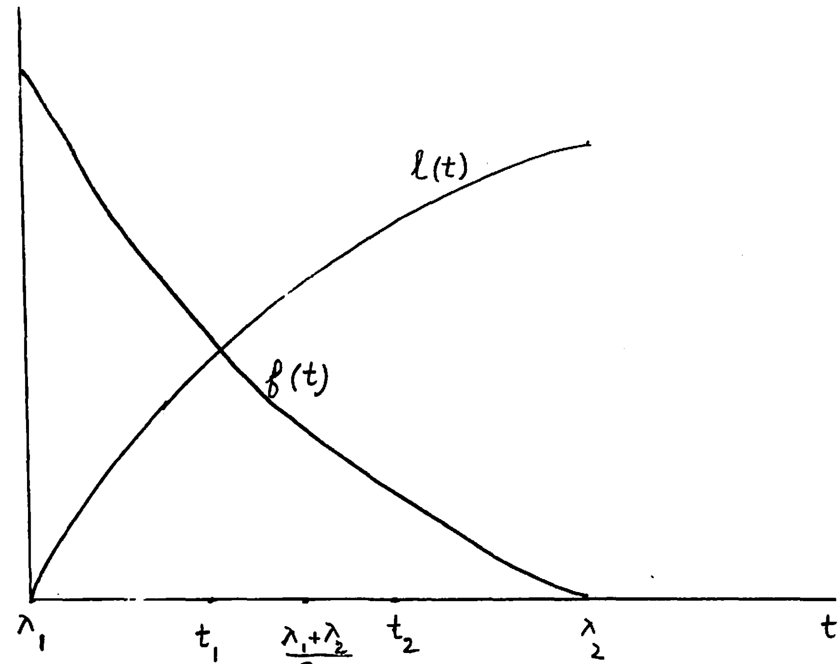

Figure 4.

Consider any points t_1 and t_2 where $t_1 = \frac{\lambda_1 + \lambda_2}{2} - t$ and $t_2 = \frac{\lambda_1 + \lambda_2}{2} + t$,

for $0 < t \leq \frac{\lambda_1 + \lambda_2}{2}$. Then,

$$L(t_1) - L(\lambda_1 + \lambda_2 - t_1) = L(\frac{\lambda_1 + \lambda_2}{2} - t) - L(\frac{\lambda_1 + \lambda_2}{2} + t). \tag{35}$$

and $$L(\lambda_1 + \lambda_2 - t_2) - L(t_2) = L(\frac{\lambda_1 + \lambda_2}{2} - t) - L(\frac{\lambda_1 + \lambda_2}{2} + t). \tag{36}$$

Thus, from (35) and (36) we have that:

$$L(t_1) - L(\lambda_1 + \lambda_2 - t_1) = L(\lambda_1 + \lambda_2 - t_2) - L(t_2). \tag{37}$$

Also, we have that $f(t_1) \geq f(t_2)$. Since L is non-decreasing we have

for $0 \leq t_1 \leq \frac{\lambda_1 + \lambda_2}{2}$, $L(t_1) - L(\lambda_1 + \lambda_2 - t_1) \leq 0$ and for

$\frac{\lambda_1 + \lambda_2}{2} \leq t_2 \leq \lambda_2$, $L(\lambda_1 + \lambda_2 - t_2) - L(t_2) \leq 0$.

$$\text{Hence,} \int_{\lambda_1}^{\frac{\lambda_1+\lambda_2}{2}} L(t_1) - L(\lambda_1 + \lambda_2 - t_1)f(t_1)dt_1 \leq$$

$$\int_{\frac{\lambda_1+\lambda_2}{2}}^{\lambda_2} \Big((L(\lambda_1+\lambda_2-t_2) - L(t_2)\Big)f(t_2)dt_2.$$

But this is equivalent to (34).

Lemma 3.4.

Assume:

(1) $f(t)$ and $L(t)$ be non-negative functions defined on $\lambda_1 \leqq t \leqq \lambda_2$.

(2) $f(t)$ is non-decreasing on (λ_1,λ_2).

(3) $L(t)$ is non-increasing on (λ_1,λ_2).

Then, $$\int_{\lambda_1}^{\lambda_2} f(t)L(t)dt \leqq \int_{\lambda_1}^{\lambda_2} f(t)L(\lambda_1+\lambda_2-t)dt. \qquad (38)$$

Proof:

The proof of this is similar to that of Lemma 3.3.

Lemma 3.5.

Assume:

(1) $f_1(t)$ is a non-negative and non-decreasing function defined on $\lambda_1 \leqq t \leqq \lambda_2$.

(2) $f_2(t)$ is a non-negative and non-increasing function defined on

$\lambda_1 + \lambda \leq b \leq \lambda_2 + \lambda$.

(3) $f_1(b) \leq f_2(\lambda_1 + \lambda_2 + \lambda - t)$, for $\lambda_1 \leq t \leq \lambda_2$.

(4) $L(t)$ is a non-positive function for $\lambda_1 \leq t \leq \lambda_2$.

Then, $$\int_{\lambda_1}^{\lambda_2} f_1(t)L(t)dt \geq \int_{\lambda_1+\lambda}^{\lambda_2+\lambda} f_2(t)L(\lambda_1 + \lambda_2 - t)dt. \tag{39}$$

Proof:

Let $\lambda_1 + \lambda_2 + \lambda - t = t'$, then

$$\int_{\lambda_1+\lambda}^{\lambda_2+\lambda} f_2(t)L(\lambda_1 + \lambda_2 + \lambda - t)dt = \int_{\lambda_1}^{\lambda_2} f_2(\lambda_1 + \lambda_2 + \lambda - t')L(t')dt' . \tag{40}$$

By hypotheses (3) and (4) of the lemma and by (40) we have that:

$$\int_{\lambda_1}^{\lambda_2} f_1(t)L(t)dt \geq \int_{\lambda_1}^{\lambda_2} f_2(\lambda_1 + \lambda_2 + \lambda - t)L(t)dt = \int_{\lambda_1+\lambda}^{\lambda_2+\lambda} f_2(t)L(\lambda_1 + \lambda_2 + \lambda - t)dt.$$

Lemma 3.6.

Let:

(1) $L(t)$ be a function defined over $-\pi \leq t < \pi$ and symmetric around 0.

(2) $L_C(t) = L_b(t) - L_{b_0}(t)$, where $L_b(t) = L(b - t)$ and $L_{b_0}(t) = L(b_0 - t)$.

Then $L_C(b + b_0 - t) = L_{b_0}(t) - L_b(t)$. $\quad (41)$

Proof:

$$\begin{aligned} L_C(b + b_0 - t) &= L_b(b + b_0 - t) - L_{b_0}(b + b_0 - t) \\ &= L(-b_0 + t) - L(-b + t) \\ &= L_{b_0}(t) - L_b(t), \qquad \text{(by symmetry of L).} \end{aligned}$$

Theorem 3.4.

Assume:

(1) $f(\alpha)$ is a circular distribution which is symmetric and unimodal.

(2) $f(\alpha)$ is non-decreasing for $b_0 - \pi \leq \alpha \leq b_0$ and non-increasing for $b_0 < \alpha < b_0 + \pi$ (so that b_0 is the mode).

(3) $L(\delta)$ is a non-negative periodic function (with period 2π) defined for $-\pi \leq \delta < \pi$.

(4) $L(\delta)$ is symmetric around 0.

(5) $L(\delta)$ is non-decreasing for $0 \leq \delta \leq \pi$.

Then,
$$\int_{b_0-\pi}^{b_0+\pi} L_{b_0}(\alpha)f(\alpha)d\alpha \leq \int_{b-\pi}^{b+\pi} L_b(\alpha)f(\alpha)d\alpha \quad , \quad \text{for all } b, \tag{42}$$
where $L_{b_0}(\alpha) = L(b_0 - \alpha)$ and $L_b(\alpha) = L(b - \alpha)$.

Proof:

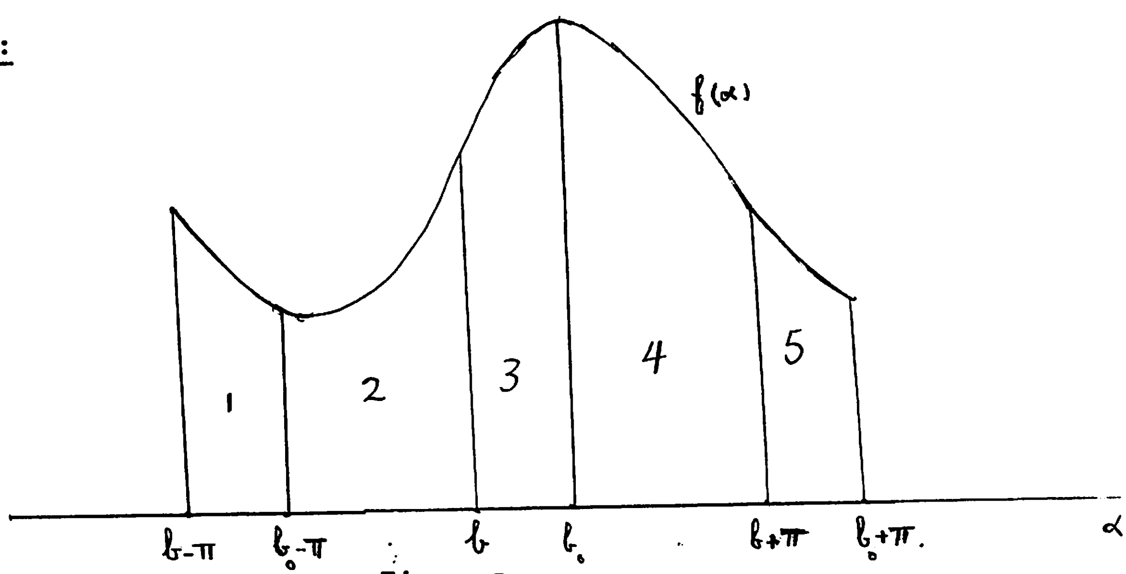

Figure 5.

Let, $b < b_0$. Then, $\int_{b-\pi}^{b+\pi} L_b(\alpha)f(\alpha)d\alpha =$

$$\int_1 L_b(\alpha)f(\alpha)d\alpha + \int_2 L_b(\alpha)f(\alpha)d\alpha + \int_3 L_b(\alpha)f(\alpha)d\alpha + \int_4 L_b(\alpha)f(\alpha)d\alpha. \tag{43}$$

$$\int_{b_0-\pi}^{b_0+\pi} L_{b_0}(\alpha)f(\alpha)d\alpha =$$

$$\int_2 L_{b_0}(\alpha)f(\alpha)d\alpha + \int_3 L_{b_0}(\alpha)f(\alpha)d\alpha + \int_4 L_{b_0}(\alpha)f(\alpha)d\alpha + \int_5 L_{b_0}(\alpha)f(\alpha)d\alpha, \quad (44)$$

where the numbers correspond to the intervals in Figure 5. By periodicity of L_{b_0} and f we have that $\int_5 L_{b_0}(\alpha)$ corresponds to the left hand side of integral of Lemma 3.3. for $\lambda_1 = b - \pi$ and $\lambda_2 = b_0 - \pi$. Also, $\int_1 L_b(\alpha)f(\alpha)d\alpha = \int_{\lambda_1}^{\lambda_2} L_{b_0}(\lambda_1 + \lambda_2 - t)f(t)dt$, with $\lambda_1 = b - \pi$ and $\lambda_2 = b_0 - \pi$. Finally, $L_{b_0}(\alpha)$ is non-decreasing over the range of integration. The hypotheses of Lemma 3.3. are thereby satisfied and we have that: $\int_5 L_{b_0}(\alpha)f(\alpha)d\alpha \leq \int_1 L_{b_0}(\lambda_\mu + \lambda_1 - t)f(t)dt = \int_1 L_b(\alpha)f(\alpha)d\alpha$. (45)

Also due to symmetry of L and with $\lambda_1 = b$ and $\lambda_2 = b_0$, we have that

$\int_3 L_b(\alpha)f(\alpha)d\alpha = \int_{\lambda_1}^{\lambda_2} L_{b_0}(\lambda_1 + \lambda_2 - t)f(t)dt$. Finally, since L_{b_0} is non-increasing, and $f(\alpha)$ is non-decreasing over the range, the hypotheses of Lemma 3.4. are satisfied. Hence, $\int_3 L_{b_0}(\alpha)f(\alpha)d\alpha \leq \int_3 L_{b_0}(\alpha)f(\alpha)d\alpha$. (46)

If we take in Lemma 3.5., the following:

(1) $f_1(t) = f(t)$, for $b_0 - \pi \leq t \leq b$,

(2) $f_2(t) = f(t)$, for $b_0 \leq t \leq b + \pi$,

(3) $\lambda_1 = b_0 - \pi$, $\lambda_2 = b$, $\lambda = \pi$,

(4) $L(t) = L_C(t) = L_b(t) - L_{b_0}(t)$, for $b_0 - \pi \leq t \leq b$.

and by observing that $L_C(t) \leq 0$ over the range $b_0 - \pi \leq t \leq b$, we see that the hypotheses of Lemma 3.5. are satisfied.

By Lemma 3.5. and since $\lambda_1 + \lambda_2 + \lambda = b_0 - \pi + b + \pi = b_0 + b$, we have that: $\int_{b_0-\pi}^{b} L_C(t)f(t)dt \geq \int_{b_0}^{b+\pi} L_C(b_0 + b - t)f(t)dt.$ (47)

But, $\int_{b_0-\pi}^{b+\pi} L_C(t)f(t)dt = \int_{b_0-\pi}^{b} (L_b(t) - L_{b_0}(t))f(t)dt$, and

$\int_{b_0}^{b+\pi} L_C(b_0 + b - t)f(t)dt = \int_{b_0}^{b+\pi} (L_{b_0}(t) - L_b(t))f(t)dt$, by (41).

Hence, $\int_{b_0-\pi}^{b} (L_b(t) - L_{b_0}(t))f(t)dt \geq \int_{b_0}^{b+\pi} (L_{b_0}(t) - L_b(t))f(t)dt$

and $\int_{2+4} L_b(t)f(t)dt \geq \int_{2+4} L_{b_0}(t)f(t)dt.$ (48)

By using (45), (46), (48), (43) and (44) we have that (42) holds. This completes the proof for the case $b < b_0$. The proof for the case $b > b_0$ is similar.

Example 3.8.

Illustration of the Theorem 3.4.

Consider again the density referred to in (6). We have the following:

(1) It is unimodal and θ is the mode.

(2) Since for any $0 \leq \gamma < 2\pi$, $f_f(\theta + \gamma) = f_f(\theta - \gamma)$, it is symmetric around θ.

(3) It is non-decreasing for $\theta - \pi \leq \alpha \leq \theta$ and non-increasing for $\theta < \alpha < \theta + \pi$.

If we take any non-negative periodic function (with period 2π) which satisfies the hypotheses of the theorem e.g. $L(\delta) = \delta^2$, then we have that:

$$\int_{\theta - \pi}^{\theta + \pi} (\theta - \alpha)^2 \, f_f(\alpha \mid r,\theta)d\alpha \;\leq\; \int_{b - \pi}^{b + \pi} (b - \alpha)^2 f_f(\alpha \mid r,\theta)d\alpha, \qquad (49)$$

for all b.

We observe that:

(1) θ, which is the mode is an invariant estimator.

(2) The left hand side of (49) is $E_f(\theta - \alpha)^2$ by (18).

(3) $E_f(\theta - \alpha)^2$ is the minimizing value of the integral on right hand side of (49) and hence $E_f(\theta - \alpha)^2 \leq E_f(\hat{\alpha} - \alpha)^2$, where $\hat{\alpha}$ is any invariant estimator. Hence, $\hat{\alpha}_{MSE}$ takes the value θ.

A more general result of this kind will be given in the following section.

3.9. Bayes' Estimate of α.

In this Section, by use of Theorem 3.4. of Section 3.8., a theorem by which a Bayes' estimate of α can be obtained in case of one or more bivariate observations is given. Certain corollaries which are special cases of the theorem are also given.

Theorem 3.5.

Assume that:

(1) the fiducial density $p(\alpha \mid \theta_1, a)$ is symmetric (Definition 3.6.) and unimodal for all θ_1, a with the mode denoted by $\theta_0 = \theta_0(\theta_1, a)$,

(2) $p(\alpha \mid \theta_1, a)$ is non-decreasing for $\theta_0 - \pi \leq \alpha \leq \theta_0$, and non-increasing for $\theta_0 < \alpha < \theta_0 + \pi$;

(3) $L(\delta) = L(\hat{\alpha} - \alpha)$, where $\hat{\alpha}$ is an estimator, is a loss function satisfying the requirements (3), (4) and (5) of Theorem 3.4. of Section 3.8.

(4) the prior density of α is uniform over the interval $(0, 2\pi)$.

Then for all θ_1, a, the Bayes' estimate of α relative to the loss function $L(\delta)$ is $\theta_0 = \theta_0(\theta_1, a)$.

Proof:

By Section 3.6.3., we have that $p(\alpha \mid \theta_1, a)$ is also the posterior density of α. Thus we want to obtain an estimator $\hat{\alpha}(\theta_1, a)$ such that $\int_{\hat{\alpha} - \pi}^{\hat{\alpha} + \pi} L(\hat{\alpha} - \alpha) p(\alpha \mid \theta_1, a)$ is minimized. By observing that all the hypotheses of Theorem 3.4. of Section 3.8. are satisfied, we have that θ_0, the mode of $p(\alpha \mid \theta_1, a)$ is the Bayes' estimate.

Corollary 3.1.

Assume that:

(1) $L(\hat{\alpha} - \alpha) = (\hat{\alpha} - \alpha)^2_m$, where $(\hat{\alpha} - \alpha)^2_m = \min\limits_{n=0,1,2,\ldots} (\hat{\alpha} - \alpha \pm 2n\pi)^2$,

(2) the hypotheses (1) and (2) of Theorem 3.5. are satisfied.

Then, θ_0, the mode of $p(\alpha \mid \theta_1, a)$, is the minimum mean square invariant estimator of α.

Proof:

It is easily seen that the $L(\hat{\alpha} - \alpha)$ given in (1) satisfies the hypothesis (3) of Theorem 3.5. Thus we have from Theorem 3.4. that: $E_f(\theta_0 - \alpha)^2 \leq E_f(b - \alpha)^2$, where b is any other estimator. We also know from Theorem 2.1. of Chapter 2 that $E_f(\hat{\alpha} - \alpha)^2 = E_R(\hat{\alpha} - \alpha)^2$, if $\hat{\alpha}$ is an invariant estimator. Since the mode is an invariant estimator, we have that $E_R(\theta_0 - \alpha)^2 \leq E_R(\hat{\alpha} - \alpha)^2$, where $\hat{\alpha}$ is any invariant estimator. It follows that $E(\theta_0 - \alpha)^2 \leq E(\hat{\alpha} - \alpha)^2$, where $\hat{\alpha}$ is any invariant estimator. Hence θ_0 is the minimum mean square invariant estimator.

Corollary 3.2.

Assume that:

(1) $L(\hat{\alpha} - \alpha) = |\hat{\alpha} - \alpha|_m$, where $|\hat{\alpha} - \alpha|_m = \min\limits_{n=0,1,2,\ldots} |\hat{\alpha} - \alpha \pm 2n\pi|$,

(2) the hypotheses (1) and (2) of Theorem 3.5. are satisfied.

Then, θ_0, the mode of $p(\alpha \mid \theta_1, a)$ is an estimator of α which has the property (27).

Proof:

It is easily verified that $L(\hat{\alpha} - \alpha)$ given in (1) satisfies the hypothesis (3) of Theorem 3.5. The rest of the proof is similar to that of Corollary 3.1.

Remark 3.4.

If the fiducial density $p(\alpha \mid \theta_1, a)$ is symmetric and unimodal; and $p(\alpha \mid \theta_1, a)$ is non-decreasing for $\theta_0 - \pi \leq \alpha \leq \theta_0$ and non-increasing for $\theta_0 < \alpha < \theta_0 + \pi$, then, θ_0, the mode of $p(\alpha \mid \theta_1, a)$ is the minimum mean square error invariant estimator of α and also possesses the property (27).

Corollary 3.3.

Assume that:

(1) $L(\hat{\alpha} - \alpha) = 1 - \operatorname{Cos}(\hat{\alpha} - \alpha)$,

(2) the hypotheses (1) and (2) of Theorem 3.5. are satisfied.

Then, θ_0, the mode of $p(\alpha|\theta_1, a)$ is a Bayes' estimate of α.

Proof:

It is easily seen that $L(\hat{\alpha} - \alpha)$ given (1) satisfies the hypotheses (3) of Theorem 3.5. The rest of the proof follows immediately from the theorem.

Chapter 4. Prediction.

4.0. Introduction.

The problem of prediction has been considered by Ramsey and Buehler (1963). They deal with the prediction of a future observation(i.e. the prediction of the (n + 1)th observation when n observations are already available). They consider the cases of families of distributions with one location parameter and one location parameter and one scale parameter. This Chapter deals with a more general problem: the prediction of certain functions of several future observations in the case of families of distributions with a group structure. Consequently, it is seen that Ramsey and Buehler's problem is a special case.

4.1. Formulation of the Problem.

The formulation of the problem is as follows. We make the following assumptions.

Assumption 1.

Let $(\underline{\overline{X}}', \beta_{\underline{\overline{X}}'}, P^{(\omega)})$ be a probability space, where $\omega \in \Omega$ and Ω is the parameter space. Also, let $\underline{\overline{X}}' = \underline{\overline{X}} \times \underline{\overline{X}}^*$, where $\underline{\overline{X}}$ is the space of "past" observations and $\underline{\overline{X}}^*$ is the space of "future" observations.

Assumption 2.

There exist spaces $\mathcal{T}, \mathcal{A}, \mathcal{A}^*$ such that $\underline{\overline{X}}'$ is in one-to-one correspondence with $\mathcal{T} \times \mathcal{A} \times \mathcal{A}^*$ and $\underline{\overline{X}}$ is in one-to-one correspondence with $\mathcal{T} \times \mathcal{A}$. For elements x', t, a, a^* of these spaces this correspondence will be denoted by:

$$x' = (x,x^*) \longleftrightarrow (t,a,a^*) \text{ and } x \longleftrightarrow (t,a). \tag{1}$$

<u>Assumption 3.</u>

Let $(\underline{\bar{X}},\beta_{\underline{\bar{X}}})$, $(\underline{\bar{X}}^*, \beta_{\underline{\bar{X}}^*})$, $(\mathcal{T}, \beta_{\mathcal{T}})$, $(\mathcal{A}, \beta_{\mathcal{A}})$, $(\mathcal{A}^*, \beta_{\mathcal{A}^*})$ and (Ω, β_Ω) be measurable spaces and let $\beta_{\mathcal{T}\times\mathcal{A}}$, $\beta_{\mathcal{T}\times\mathcal{A}^*}$, $\beta_{\mathcal{A}\times\mathcal{A}^*}$ and $\beta_{\mathcal{T}\times\mathcal{A}\times\mathcal{A}^*}$ be the minimal σ-fields which contain the cartesian products of the sets in the given σ-fields. Also, for $T\epsilon\beta_{\mathcal{T}}$, $A\epsilon\beta_{\mathcal{A}}$ and $A^*\epsilon\beta_{\mathcal{A}^*}$, if:

$$X' = \{x' : x' \longleftrightarrow (t,a,a^*) \text{ for } t\epsilon T,\ a\epsilon A \text{ and } a^*\epsilon A^*\}, \tag{2(a)}$$

then $X'\epsilon\beta_{\underline{\bar{X}}'}$.

Finally, for $X'\epsilon\beta_{\underline{\bar{X}}'}$, we have that if:

$$B' = \{(t,a,a^*) : (t,a,a^*) \longleftrightarrow x' \text{ for } x'\epsilon X'\}, \tag{2(b)}$$

then $B'\epsilon\beta_{\mathcal{T}\times\mathcal{A}\times\mathcal{A}^*}$.

Since $x' \longleftrightarrow (t,a,a^*)$, (t,a^*) is conditionally sufficient for ω given a. Also, the measures P^ω on $\underline{\bar{X}}$ impose corresponding measures on $\mathcal{T}\times\mathcal{A}\times\mathcal{A}^*$, i.e. for every $T\epsilon\beta_{\mathcal{T}}$, $A\epsilon\beta_{\mathcal{A}}$, $A^*\epsilon\beta_{\mathcal{A}^*}$,

$$P^\omega(T \times A \times A^*) = P^\omega(X'), \tag{3}$$

where X' and $T \times A \times A^*$ are related through (2(a)).

<u>Remark 4.1.</u>

There is no danger of confusion in using the same symbol P^ω for measures on spaces $\underline{\bar{X}}'$ and $\mathcal{T}\times\mathcal{A}\times\mathcal{A}^*$.

<u>Assumption 4.</u>

There is a group $\mathcal{G} = \{g\}$ of one-to-one measurable transformations on the sample space $\underline{\bar{X}}'$ onto itself and $(\mathcal{G}, \beta_{\mathcal{G}})$ is a measurable space.

Also assume that there exists a left Haar measure μ on the space having the invariance property given by:

$$\mu(gG) = \mu(G), \tag{4}$$

for all $g \epsilon \mathcal{G}$ and $G \epsilon \mathcal{B}_{\mathcal{G}}$.

Assumption 5.

The class of measures P^{ω} is closed under $\mathcal{G}$. Thus for any $g \epsilon \mathcal{G}$ and $\omega \epsilon \Omega$, there is a $\omega_g \epsilon \Omega$ such that for all $X' \epsilon \mathcal{B}_{\underline{\bar{X}}'}$,

$$P^{\omega}(X') = P^{g^*\omega}(gX'), \tag{5}$$

where $g^*\omega = \omega_g$.

The transformations $\mathcal{G}^* = \{g^*\}$ form a group.

Assumption 6.

For any ω_1, $\omega_2 \epsilon \Omega$, there is a single $g \epsilon \mathcal{G}$ such that $g^*\omega_1 = \omega_2$.

That is, $\mathcal{G}^*$ is exactly transitive on Ω.

Assumption 7.

For any $g \epsilon \mathcal{G}$ and $x' \epsilon \underline{\bar{X}}'$, if $x' \longleftrightarrow (t,a,a^*)$, assume that

$$gx' \longleftrightarrow (t_g,a,a^*), \tag{6}$$

where t_g depends only on t and g and not a or a^*.

Definition 4.1.

Let g' be defined by:

$$g't = t_{g} . \tag{7}$$

It is seen that $\mathcal{G}' = \{g'\}$ is a group which is isomorphic to $\mathcal{G}$.

Assumption 8.

$\mathcal{G}'$ is exactly transitive on $\mathcal{T}$.

Representation of Spaces $\mathcal{T}$ and Ω respectively in terms of $\mathcal{G}'$ and $\mathcal{G}^*$.

Let x_0' and ω_0 be arbitrary but fixed reference points in the sample space $\underline{\bar{X}}'$ and parameter space Ω. If $x_0' \longleftrightarrow (t_0, a_0, a_0^*)$, than t_0, a_0, and a_0^* are taken to be the corresponding reference points in spaces $\mathcal{T}$, $\mathcal{A}$ and $\mathcal{A}^*$ respectively. Let g_t' be the unique transformation in $\mathcal{G}'$ which for each a, a^* carries t_0 into t and g_ω' be the transformation which carries the conditional variable t_0 with ω_0 distribution into a conditional variable t with ω distribution. Also, let g_ω^* be the transformation that carries ω_0 into ω. We will denote g_t' and g_ω' in $\mathcal{G}'$ by t and ω respectively and g_ω^* in $\mathcal{G}^*$ by ω^*.

Remark 4.2.

The correspondence among the elements of transformations $\mathcal{G}$, $\mathcal{G}'$ and $\mathcal{G}^*$ is obvious.

Assumption 9.

Let $\underline{\bar{X}} \subseteq R_{n'}$, (the n'-dimensional Euclidean space). Assume that for all $\omega \epsilon \Omega$, the measure P^ω is absolutely continuous with respect to the n'-dimensional Lebesgue measure $L_{n'}$, i.e., $(P^\omega << L_{n'})$.

Then, by Radon-Nikodyn Theorem, there exists a $L_{n'}$-measurable function $p(x',\omega)$ such that :

$$P^{\omega}(X') = \int_{X'} p(x',\omega)dL_{n'}(x'), \tag{8}$$

for every $X'\epsilon B_{\underline{X}'}$.

Lemma 4.1.

If (5) holds, then for all g,ω:

$$p(x',\omega) = p(gx',g^{*}\omega), \text{ a.e. } L_{n'}, \tag{9}$$

where $g \longleftrightarrow g^{*}$.

Proof:

The proof is similar to that of Lemma 2.1.

Lemma 4.2.

If for all x',g,ω (where $x' \longleftrightarrow (t,a,a^{*})$ and $g \longleftrightarrow g' \longleftrightarrow g^{*}$),

$$H(x',\omega) = H(gx',g^{*}\omega), \tag{10}$$

then $H(x',\omega)$ can be expressed in the form $H'(\omega^{-1}t,a,a^{*})$.

Proof:

The proof of this Lemma is similar to that of Lemma 2.2.

Assumption 10.

Assume there exists a measure λ_1 on $\beta_{\mathcal{A}}$ and λ_2 on $\beta_{\mathcal{A}^{*}}$ and a function $h(t,a,a^{*})$ such that for $S\epsilon B_{n'}$ (the n'-dimensional Borel field), we have that:

$$\int_{S} dL_{n'}(x') = \int_{S'} h(t,a,a^{*})d\mu(t)d\lambda_1(a)d\lambda_2(a^{*}), \tag{11}$$

where $S' = \{(t,a,a^{*}) : x' \longleftrightarrow (t,a,a^{*}) \text{ if } x'\epsilon S\}$.

Definition 4.2.

For $A \times A^* \epsilon \mathcal{B}_{\mathcal{A} \times \mathcal{A}^*}$, define:

$$P_2(A \times A^*) \;=\; P^{\omega}(\mathcal{T} \times A \times A^*) \tag{12}$$

and for $A \epsilon \mathcal{B}_{\mathcal{A}}$ define:

$$P_3^{\omega}(A) \;=\; P^{\omega}(\mathcal{T} \times A \times \mathcal{A}^*). \tag{13}$$

Lemma 4.3.

$$P_2^{\omega}(\cdot) \;=\; P_2(\cdot) \text{ and } P_3^{\omega}(\cdot) \;=\; P_3(\cdot).$$

Proof:

For $T \times A \times A^* \epsilon \mathcal{B}_{\mathcal{T} \times \mathcal{A} \times \mathcal{A}^*}$, we have that:

$$P^{\omega}(T \times A \times A^*) \;=\; \int_C dP^{\omega}(x') = \int_D dP^{g^*\omega}(gx') = P^{g^*\omega}(T_g \times A \times A^*), \tag{14}$$

where $C \;=\; \{x' : x' \longleftrightarrow (t,a,a^*) \;\epsilon T \times A \times A^*\}$,

$D \;=\; \{gx' : gx' \longleftrightarrow (t_g,a,a^*) \;\epsilon T_g \times A \times A^*\}$

and $T_g \times A \times A^* \;=\; \{(t_g,a,a^*)$ if $x' \longleftrightarrow (t,a,a^*) \;\epsilon T \times A \times A^*$, then
$gx' \longleftrightarrow (t_g,a,a^*)\}$.

Also by (12), (13) and since $\mathcal{T}_g = \mathcal{T}$, we have that:

$$P_2^{\omega}(A \times A^*) = P^{\omega}(\mathcal{T} \times A \times A^*) = P^{g^*\omega}(\mathcal{T}_g \times A \times A^*) = P^{g^*\omega}(\mathcal{T} \times A \times A^*)$$
$$= P_2^{g^*\omega}(A \times A^*).$$

Since this is true for all $g^*\omega$ and $\{g^*\omega : g^* \epsilon \mathcal{G}^*\} = \{\omega : \omega \epsilon \Omega\}$, it follows that $P_2^{\omega}(\cdot) = P_2(\cdot)$. Also, since from (13) and (12) we have that:
$P_3^{\omega}(A) = P^{\omega}(\mathcal{T} \times A \times \mathcal{A}^*) = P_2^{\omega}(A \times \mathcal{A}^*) = P_2(A \times \mathcal{A}^*)$. Hence $P_3(\cdot)$ does

not depend upon ω. This completes the proof of the lemma.

Due to (13), we have that P^{ω} considered as a measure on space $\mathcal{A}$ (for fixed $T \times A^*$) is absolutely continuous with respect to P_3. Hence we can define a $\beta_{\mathcal{A}}$-measurable function $P_1^{\omega}(T \times A^* \mid a)$ by use of the Randon-Nikodyn Theorem according to:

$$P^{\omega}(T \times A \times A^*) = \int_A P_1^{\omega}(T \times A^* \mid a) dP_3(a), \tag{15}$$

for all $A \epsilon \beta_{\mathcal{A}}$.

Remark 4.3.

For almost all a, $P_1(\cdot \mid a)$ is a conditional probability measure on space $\mathcal{T} \times \mathcal{A}^*$.

Assumption 11.

Assume that for all a and ω, $P_1^{\omega} << \mu \times \lambda_2$. By Assumption 11 and application of Radon-Nikodyn Theorem, we can define a $\beta_{\mathcal{T}}$-measurable function $p_1(t,a^* \mid a;\omega)$ by:

$$P_1^{\omega}(T \quad A^* \mid a) = \int_{T \times A^*} p_1(t,a^* \mid a,\omega) d\mu(t) d\lambda_2(a^*), \tag{16}$$

where $p_1(t,a^* \mid a,\omega)$ is a $\beta_{\mathcal{T} \times \mathcal{A}^*}$-measurable function.

Remark 4.4.

$p_1(t,a^* \mid a;\omega)$ is a conditional density on space $\mathcal{T} \times \mathcal{A}^*$ with respect to the measure $\mu \times \lambda_2$.

Lemma 4.4.

$$p_1(t,a^* \mid a,\omega) = p_1'(\omega^{-1}t,a^* \mid a). \tag{17}$$

By (15) we have that:

$$P^{\omega}(T \times A \times A^{*}) = \int_{A} P_1^{\omega}(T \times A \mid a)dP_3(a), \text{ for all } A\epsilon\beta_{\mathcal{A}} .$$

Similarly, letting $T_g = \{t_g : t\epsilon T\}$ we have that:

$$P^{g^{*}\omega}(T_g \times A \times A^{*}) = \int_{A} P_1^{g^{*}\omega}(T_g \times A^{*} \mid a)dP_3(a).$$

But by (3),

$$P^{\omega}(T \times A \times A^{*}) = P^{g^{*}\omega}(T_g \times A \times A^{*}).$$

Hence, $P_1^{\omega}(T \times A^{*} \mid a) = P_1^{g^{*}\omega}(T_g \times A^{*} \mid a)$, a.e, P_3.

Also by (16) we have that:

$$P_1^{\omega}(T \times A^{*} \mid a) = \int_{T \times A^{*}} p_1(t,a^{*} \mid a,\omega)d\mu(t)d\lambda_2(a^{*})$$

and $P_1^{g^{*}\omega}(T_g \times A^{*} \mid a) = \int_{T_g \times A^{*}} p_1(t,a^{*} \mid a,g^{*}\omega)d\mu(t)d\lambda_2(a^{*})$

$$= \int_{T \times A^{*}} p_1(t_g,a^{*} \mid a,g^{*}\omega)d\mu(t_g)d\lambda_2(a^{*})$$

$$= \int_{T \times A^{*}} p_1(t_g,a^{*} \mid a,g^{*}\omega)d\mu(t)d\lambda_2(a^{*}), \text{ by (14)}.$$

Hence, $p_1(t,a^{*} \mid a,\omega) = p_1(t_g,a^{*} \mid a,g^{*}\omega)$, a.e. $\mu \times \lambda_2$.

Put $g^{*} = \omega^{-1}$ and we have that (19) holds.

<u>Definition 4.3.</u>

Define a measure ν on space $\mathcal{T}$ (which imposes corresponding measures on spaces $\mathcal{G}'$ and $\mathcal{G}$) as follows:

$$\nu(T) = \mu(T^{-1}), \tag{18}$$

for $T\epsilon\beta_{\mathcal{T}}$.

Remark 4.5.

ν is a right Haar measure.

Remark 4.6.

There exists a modular function Δ such that:

$$\mu(Tg') = \Delta(g')\mu(T), \tag{19}$$

for all $T \in \mathcal{B}_{\mathcal{T}}$. This is due to the fact that invariant measures are unique up to a constant and $\mu_{g'}$; where $\mu_{g'}(T) = \mu(Tg')$ is another left Haar measure.

Remark 4.7.

The measure elements of ν and μ are related as follows:

$$d\mu(t) = \Delta(t)d\nu(t). \tag{20}$$

Assumption 12.

Let F be a transformation from $\mathcal{T} \times \mathcal{A} \times \mathcal{A}^*$ onto $\underline{\bar{X}} \times \underline{\bar{X}}^*$, such that:

$$F((t,a),a^*) = (x,x^*), \tag{21}$$

where x <----> (t,a).

Let $F_{(t,a)}$ denote the transformation which corresponds to the second argument of F for fixed (t,a) (i.e., for fixed x). Thus it is a mapping of $\mathcal{A}^*$ onto $\underline{\bar{X}}^*$, for fixed x. Also for $X^* \in \mathcal{B}_{\underline{\bar{X}}^*}$, $F^{-1}_{(t,a)}(X^*) \in \mathcal{B}_{\mathcal{A}^*}$.

Finally, assume that $\lambda_2 F^{-1}_{(t,a)} << L_{n'-n}$.

By Theorem D on page 164 in Halmos (1950) we have that for $A^* \in \mathcal{B}_{\mathcal{A}^*}$ and $X^* \in \mathcal{B}_{\underline{\bar{X}}^*}$, there exists a $\mathcal{B}_{\underline{\bar{X}}^*}$ -measurable function $\varphi(x,x^*)$ such that:

$$\int_{A^*} d\lambda_2(a^*) = \int_{X^*} \varphi(x,x^*)dL_{n'-n}(x^*), \tag{22}$$

where $X^* = \{x^* : (x,x^*) \longleftrightarrow (t,a,a^*) \text{ for } a^* \epsilon A^*\}$. (23)

Fisher (1935, 1956) considered for the normal distribution case: (i) the fiducial density of an independent future observation, and (ii) the joint fiducial density of sample statistics $\bar{x}$ and s ($\bar{x}$ = sample mean, and s = sample standard deviation) from a subsequent sample. Fraser (1961) has defined the fiducial density of ω as pointed out in Chapter 2. More recently Ramsey and Buehler (1963) have defined the joint density of ω and x^*, when $x^* = x_{n+1}$. The definition which we give below of the joint fiducial density of ω and x^* is partly on lines similar to those of Fraser (1961) and agrees with Ramsey and Buehler when $x^* = x_{n+1}$.

Definition 4.4.

Define the joint fiducial density of ω and x^* with respect to the measure $\nu \times L_{n'-n}$ as follows:

$$p(\omega,x^* \mid x) = p_1'(\omega^{-1}t,a^* \mid a)\,\Delta(t)\varphi(x,x^*). \tag{24}$$

4.2. Expectation Identity.

In this Section we obtain the expectation identity, $E_R(H(x',\omega)) = E_f(H(x',\omega))$, where $H(x',\omega)$ is a function for which (10) holds. We begin with the theorem which proves the above mentioned expectation identity.

Theorem 4.1.

Let E_R denote the conditional expecatation with respect to the density $p_1(t,a^* \mid a,\omega)$ defined in (16) and E_f denote the conditional expectation with respect to the fiducial density given by (24).

Assume:

(1) Assumptions 1-12 are satisfied.

(2) $H(x',\omega)$ is an invariant function i.e. (10) holds. Then:

$$E_R(H(x',\omega)) = E_f(H(x',\omega)). \tag{25}$$

Proof:

$$E_R(H(x',\omega)) = \int_{t\in T,\, a^*\in \mathcal{A}^*} H'(\omega^{-1}t,a,a^*)p_1(t,a^* \mid a,\omega)d\mu(t)d\lambda_2(a^*),$$

by Lemma 4.2.

$$= \int_{t\in T,\, a\in \mathcal{A}^*} H'(\omega^{-1}t,a,a^*)p_1'(\omega^{-1}t,a^* \mid a)d\mu(t)d\lambda_2(a^*), \text{ by (17).}$$

$$= \int_{s\in\omega^{-1}T=T,\, a^*\in\mathcal{A}^*} H_*'(s,a,a^*)p_1'(s,a^*,a \mid a)d\mu(s)d\lambda_2(a^*)$$

$$= \int_{\omega\in\Omega,\, a^*\in \mathcal{A}^*} H'(\omega^{-1}t,a,a^*)p_1'(\omega^{-1}t,a^* \mid a)\Delta(t)d\mu(\omega^{-1})d\lambda_2(a^*),$$

by (19), for a fixed t.

$$= \int_{\omega\in\Omega,\, a^*\in \mathcal{A}^*} H'(\omega^{-1}t,a,a^*)p_1'(\omega^{-1}t,a^* \mid a)\Delta(t)d\nu(\omega)d\lambda_2(a^*),$$

by (18).

$$= \int_{\omega\in\Omega, x^*\in\overline{\underline{X}}^*} H'(\omega^{-1}t,a,a^*)p_1'(\omega^{-1}t,a^* \mid a)\Delta(t)\varphi(x,x^*)dL_{n'-n}(x^*)d\nu(\omega),$$

by (22).

$$= E_f(H(x',\omega)), \text{ by } (24).$$

In the second last line it is to be understood that (t,a,a^*) corresponds to (x,x^*) through (1).

4.3. Case of One Location Parameter.

In this Section we consider a theorem which deals with families of distributions with one location parameter θ.

Definition 4.5.

A function $B(x) = B(x_1, \ldots, x_n)$ is said to be a "location invariant" function if for any a, $-\infty < a < \infty$,

$$B(x_1, \ldots, x_n) - a = B(x_1 - a, \ldots, x_n - a). \tag{26}$$

Definition 4.6.

For $\underline{\bar{X}}' = R_{n'}$, define g on $\underline{\bar{X}}'$ as follows:

$$gx' = (x_1 - g, \ldots, x_{n'} - g). \tag{27}$$

For $\Omega = R_1$, define g^* on $\Omega = \{\theta\}$ as follows:

$$g^*\theta = \theta - g. \tag{28}$$

Ramsey and Buehler (1963) have obtained the expectation identity for functions of the form $\varphi(B(x) - x_{n+1})$. With the above definitions of g and g^* and with $\theta = \omega$, it is easily seen that these functions satisfy the invariance relation (10) for $H(x',\omega)$. Thus the following theorem which will be shown to follow from Theorem 4.1., is very similar to Ramsey and Buehler's result. It is actually more general for the function $H(x',\omega)$ is more general.

Theorem 4.2.

Let E_R denote the conditional expectation over a region R in which the ancillaries $a_i = x_{i+1} - x_1$ $(i = 1, \ldots, n-1)$ are fixed and E_f denote the expectation with respect to the fiducial density of θ and x^* given x.

Assume:

(1) $x' = (x, x^*)$ has density with respect to the Lebesgue measure $L_{n'}$ given by:

$$p(x',\theta) = f(x_1 - \theta, \ldots, x_{n'} - \theta), \; -\infty < x_1, \ldots, x_{n'} < \infty, \tag{29}$$

where $x = (x_1, \ldots, x_n)$, $x^* = (x_{n+1}, \ldots, x_{n'})$ and $n' > n$.

(2) $H(x',\theta)$ is an invariant function i.e., for which (10) holds.

Then

$$E_R(H(x',\theta)) = E_f(H(x',\theta)). \tag{30}$$

<u>Proof</u>:

We begin with the verification of Assumptions of Theorem 4.1. Since this verification is very similar to that of Theorem 2.2., it will be gone into briefly.

We have the following:

(1) For $X' \in \beta_{\bar{\underline{X}}'}$, we have that:

$$p^{\omega}(X') = \int_{X'} f(x_1 - \theta, \ldots, x_{n'} - \theta) dx_1 \ldots dx_{n'} . \tag{31}$$

(2) In the correspondence $x' \longleftrightarrow (t,a,a^*)$, where $x \longleftrightarrow (t,a)$, t,a,a^* are given by:

$t = x_1$, $a = (a_1, \ldots, a_{n-1})$ $a^* = (a_n, \ldots, a_{n'-1})$, where $a_i = x_{i+1} - x_1 (i = 1, \ldots, n'-1)$.

(3) g is defined on $\bar{\underline{X}}'$ as in (26) and g^* on Ω as in (27).

(4) In Assumption 10, we have that:

$$d\lambda_1(a) = da_1 \ldots da_{n-1}, \; d\lambda_2(a^*) = da_n \ldots da_{n'-1} \tag{32}$$

and $h(t,a,a^*) = 1$.

(5) Since for $X^* \in \beta_{\bar{\underline{X}}^*}$,

$$F^{-1}_{(t,a)}(X^*) = \{a^* : a^* = (x_{n+1} - x_1, \ldots, x_{n'} - x_1)$$
$$\text{for } x^* = (x_{n+1}, \ldots, x_{n'}) \epsilon X^*\},$$

we have that $F^{-1}_{(t,a)}(X^*) \epsilon B_{n'-n}$.

Also for any $A^* \epsilon \beta_{\mathcal{A}^*}$ and $X^* \epsilon \beta_{\underline{X}^*}$, we have that:

$$\int_{A^*} da^*_{n+1} \ldots da_{n'-1} = \int_{X^*} dx_{n+1} \ldots dx_{n'} ,$$

where X^* and A^* are related through (23).

Thus $\varphi(x,x^*) = 1$.

(6) The conditional density $p_1(t,a^* \mid a,\omega)$ with respect to the measure $\mu \times \lambda_2$ (where $d\mu(t) = dt$ and $d\lambda_2$ is given in (32)) is given by:

$$p_1(t,a^* \mid a,\omega) = \frac{f(t - \theta, a_1 + t - \theta, \ldots, a_{n'-1} + t - \theta)}{\int_T \int_{\mathcal{A}^*} f(t - \theta, a_1 + t - \theta, \ldots, a_{n'-1} + t - \theta) d\mu(t) d\lambda_2(a^*)} \tag{33}$$

The denominator equals the marginal density of a and therefore the integral exists.

(7) Finally we have that:

$$\Delta(t) = 1 \text{ and } d\nu(\theta) = d\theta.$$

Also, the fiducial density of θ and x^* given x by use of (24) is given by:

$$p(\theta,x^* \mid x) = p_1(t,a^* \mid a,\omega). \tag{34}$$

Thus we see that all the hypotheses of Theorem 4.1. are satisfied and consequently (30) holds.

For the proofs of the following Corollaries, the reader is referred to Ramsey and Buehler (1963).

Corollary 4.1.

If Assumption (1) of Theorem 4.2. is satisfied, $x = (x_1, \ldots, x_n)$ and $x^* = x_{n+1}$, then the fiducial mean B_M is the minimum mean square error invariant predictor of x_{n+1}.

Corollary 4.2.

If Assumption (1) of Theorem 4.2. is satisfied, $x = (x_1, \ldots, x_n)$ and $x^* = x_{n+1}$, then the fiducial median B_C satisfies:

$$E(|B_C - x_{n+1}|) \leq E(|B - x_{n+1}|), \tag{35}$$

where B is any invariant predictor.

Definition 4.7.

An invariant predictor B_b is said to be the "best" predictor of x_{n+1} (in the Pitman sense) if for all $h > 0$,

$$P(|B_b - x_{n+1}| \leq h) \geq P(|B - x_{n+1}| \leq h) \tag{36(a)}$$

and for some $h > 0$

$$P(|B_b - x_{n+1}| \leq h) > P(|B - x_{n+1}| \leq h). \tag{36(b)}$$

Corollary 4.3.

If the fiducial density $p(x_{n+1} \mid x)$ of x_{n+1} given x:

(1) has a mode at B_L, say,

(2) is symmetric

(3) satisfies the (unimodal) conditions:

$$p(x'_{n+1} \mid x) \leq p(x''_{n+1} \mid x), \text{ for } -\infty \leq x'_{n+1} \leq x''_{n+1} \leq B_L$$

and $$p(x'_{n+1} \mid x) \geq p(x''_{n+1} \mid x), \text{ for } B_L < x'_{n+1} \leq x''_{n+1} < \infty .$$

Then $B_M = B_C = B_L (= B_b)$.

4.4. Case of One Location Parameter and One Scale Parameter.

In this Section we consider a theorem which deals with families of distributions with one location parameter θ and one scale parameter σ.

Definition 4.8.

A function $B(x) = B(x_1, \ldots, x_n)$ is said to be invariant function if for any a, $-\infty < a < \infty$, $b > 0$,

$$b[B(x_1, \ldots, x_n) - a] = B[b(x_1 - a), \ldots, b(x_n - a)]. \tag{37}$$

Definition 4.9.

Let $\underline{\overline{X}}' = R_{n'}$ define g on $\underline{\overline{X}}'$ as follows:

$$g_{a,b}\, x' = (b(x_1 - a), \ldots, b(x_{n'} - a)). \tag{38}$$

For $\Omega = \{(a,b), -\infty < a < \infty, 0 < b < \infty\}$ define $g^*_{a,b}$ as follows:

$$g^*_{a,b}\,(\theta,\sigma) = (b(\theta - a), b\sigma). \tag{39}$$

Ramsey and Buehler (1963) have obtained the expectation identity for functions of the form $\varphi\left(\frac{B(x_1, \ldots, x_n) - x_{n+1}}{\sigma}\right)$. With the above definitions of g and g^* and with $\omega = (\theta,\sigma)$ it is easily seen that the functions satisfy the invariance relation (10) for $H(x',\omega)$. Thus, the following theorem which will be shown to follow from Thorem 4.1. is very similar to Ramsey and Buehler's result. It is actually more general for the function $H(x',\omega)$ is more general.

Theorem 4.3.

Let $a_0 = 1$ or 0, according as $x_2 \geqq x_1$ or $x_2 < x_1$ and let $a_i = \frac{x_{i+2} - x_1}{x_2 - x_1}$ $(i = 1, \ldots, n-2)$. Also let E_R denote the conditional expectation over a region R in which the ancillaries $a_0, a_1, \ldots, a_{n-2}$ are fixed and E_f denote the expectation with respect to the fiducial density of θ,σ and x^* given x.

Assume:

(1) $x' = (x,x^*)$ has the density given by:

$$p(x',\theta,\sigma) = \sigma^{-n'} f\left(\frac{x_1 - \theta}{\sigma}, \ldots, \frac{x_{n'} - \theta}{\sigma}\right), \tag{40}$$

$-\infty < x_1, \ldots, x_{n'}, \theta < \infty$, $\sigma > 0$, where $x = (x_1, \ldots, x_n)$, $x^* = (x_{n+1}, \ldots, x_{n'})$ and $n' > n$. Then:

$$E_R(H(x',\theta,\sigma)) = E_f(H(x',\theta,\sigma)). \tag{41}$$

<u>Proof:</u>

Again the verification of the assumptions will be gone into briefly.

We have the following:

(1) In the correspondence $x' \longleftrightarrow (t,a,a^*)$, t, a, a^* are given by:

$t = (t_1,t_2) = (x_1, |x_2 - x_1|)$, $a = (a_0, \ldots, a_{n-2})$, $a^* = (a_{n-1}, \ldots, a_{n'-2})$,

where $a_j = \dfrac{x_{j+2} - x_1}{x_2 - x_1}$, for $j = n-1, \ldots, n'-2$.

(2) The definitions of $g_{a,b}$ and $g^*_{a,b}$ are given in (38) and (39). The measure element of the measure μ is given by $\dfrac{da\,db}{b^2}$.

(3) The measure elements $d\mu(t)$, $d\lambda_1(a)$ and $d\lambda_2(a^*)$ are given by:

$$d\mu(t) = \frac{dt_1 dt_2}{t_2^2},$$

$$d\lambda_1(a) = da_1 \ldots da_{n-2}\, d\eta, \tag{42}$$

$$d\lambda_2(a^*) = da_{n-1} \ldots da_{n'-2},$$

where η is a discrete measure having mass $\frac{1}{2}$ at each of the two values $a_0 = 0, 1$.

(4) It is easily seen that for sets $S' \epsilon \beta_{T \times \mathcal{A} \times \mathcal{A}^*}$ and $S \epsilon \beta_{\underline{X}'}$, where S and S' are related through (2(b)), we have that:

$$\int_{S'} 2t_2^{n'}\, d\mu(t) d\lambda_1(a) d\lambda_2(a^*) = \int_S dL_{n'}(x'). \tag{43}$$

Hence, $h(t,a,a^*) = 2t_2^{n'}$.

(5) For any $A^* \epsilon \beta_{\mathcal{A}^*}$ and $X^* \epsilon \beta_{\underline{X}^*}$ we have that:

$$\int_{A^*} da_{n-1} \ldots da_{n'-2} = \int_{X^*} |x_2 - x_1|^{-(n'-n)} dx_{n+1} \ldots dx_{n'}, \tag{44}$$

where X^* and A^* are related through (23).

Thus $\varphi(x,x^*) = |x_2 - x_1|^{-(n'-n)}$.

(6) The donditional density $p_1(t,a^* \mid a,\omega)$ with respect to the measure $\mu \times \lambda_2$, where $d\mu(t)$ and $d\lambda_2(a^*)$ are as in (42) is given by:

$$p_1(t,a^* \mid a,\omega) = \frac{t_2^{n'} f(\frac{x_1 - \theta}{\sigma}, \ldots, \frac{x_{n'} - \theta}{\sigma})}{\int_T \int_{\mathcal{A}^*} t_2^{n'} f(\frac{x_1 - \theta}{\sigma}, \ldots, \frac{x_{n'} - \theta}{\sigma}) d\mu(t) d\lambda_2(a^*)}. \tag{45}$$

(7) Finally we have that:

$$\Delta(t) = \frac{1}{t_2} \text{ and } d\nu(\omega) = \frac{d\theta d\sigma}{\sigma}.$$

Also, the fiducial density of θ,σ,x^* with respect to the measure $\nu \times \lambda_2$ is given by:

$$p(\theta,\sigma,x^* \mid x) = p_1(t,a^* \mid a,\theta,\sigma)\ |x_2 - x_1|^{-(n'-n+1)}. \tag{46}$$

Thus we see that all the hypotheses of Theorem 4.1. are satisfied and we have that (41) holds.

For the proof of the following corollaries, the reader is referred to Ramsey and Buehler (1963).

<u>Corollary 4.4.</u>

If Assumption (1) of Theorem 4.3. is satisfied, $x = (x_1, \ldots, x_n)$ and $x^* = x_{n+1}$, then $E_f(\frac{x_{n+1}}{\sigma^2}) \,/\, E_f(\frac{1}{\sigma^2})$ is the minimum mean square error invariant predictor of x_{n+1}.

<u>Definition 4.10.</u>

A function $B(x_1, \ldots, x_n)$ is said to be invariant if for all a,b, $-\infty < a < \infty$, $b > 0$, we have that:

$$b(B(x_1, \ldots, x_n)) = B(b(x_1 - a), \ldots, b(x_n - a)). \tag{47}$$

<u>Corollary 4.5.</u>

Let $\bar{x}^* = \frac{1}{n'-n} \sum_{j=n+1}^{n'} x_j$ and $s^2 = \frac{1}{n'-n-1} \sum_{j=n+1}^{n'} (x_j - \bar{x}^*)^2$. If Assumption (1) of Theorem 4.3. is satisfied, $x = (x_1, \ldots, x_n)$ and $x^* = (x_{n+1}, \ldots, x_{n'})$, $n' \geq n + 2$, then the minimum mean square error invariant predictor of s in the future sample (i.e. $x_{n+1}, \ldots, x_{n'}$) is given by: $E_f(\frac{s}{\sigma^2}) \,/\, E_f(\frac{1}{\sigma^2})$.

Proof:

It is easily seen that $\left(\frac{B(x_1, \ldots, x_n) - s}{\sigma}\right)^2$ satisfies (10), where $B(x_1, \ldots, x_n)$ is an invariant function according to (47). We want to obtain a $B(x_1, \ldots, x_n)$ such that $E(B(x_1, \ldots, x_n) - s)^2$ is minimum or equivalently we want to determine a $B(x_1, \ldots, x_n)$ such that $E\left[\frac{B(x_1, \ldots, x_n) - s}{\sigma}\right]^2$ is minimized. By Theorem 4.3., we have that:

$$E_R\left[\frac{B(x_1, \ldots, x_n) - s}{\sigma}\right]^2 = E_f\left[\frac{B(x_1, \ldots, x_n) - s}{\sigma}\right]^2. \qquad (48)$$

Also if the right-hand side of (48) is minimum implies

$0 = E_f\left[\frac{B(x_1, \ldots, x_n) - s}{\sigma^2}\right] = B(x_1, \ldots, x_n)\, E_f\left(\frac{1}{\sigma^2}\right) - E_f\left(\frac{s}{\sigma^2}\right)$. Hence,

$B(x_1, \ldots, x_n) = E_f\left(\frac{s}{\sigma^2}\right) / E_f\left(\frac{1}{\sigma^2}\right)$. Consequently, it is the minimum mean square error invariant predictor of s.

Definition 4.11.

A function $B(x_1, \ldots, x_n)$ will be said to be an invariant predictor if for all a,b, $-\infty < a < \infty$ and $b > 0$, we have that:

$$b^2(B(x_1, \ldots, x_n)) = B(b(x_1 - a), \ldots, b(x_n - a)). \qquad (49)$$

Remark 4.8.

Similarly it can be shown that the minimum mean square error invariant predictor of s^2 is $E_f\left(\frac{s^2}{\sigma^4}\right) / E_f\left(\frac{1}{\sigma^4}\right)$, where invariant predictor is defined according to (49).

Remark 4.9.

If any function $S(x_{n+1}, \ldots, x_{n'})$ of the future observations has the invariance property according to (47), then $E_f(\frac{S}{\sigma^2}) \,/\, E_f(\frac{1}{\sigma^2})$ is the minimum mean square error invariant predictor of S.

4.5. Case of Two Location Parameters and One Scale Parameter.

In this Section we consider a theorem which deals with families of distributions with two location parameters and one scale parameter. A feature of this theorem is that by its application we will be able to obtain an invariant predictor for the difference of two future observations i.e. $x_{m+1} - y_{n+1}$.

Definition 4.12.

For $\underline{\overline{X}}' = R_{m'+n'}$ define $g_{a_1,a_2,b}$ on $\underline{\overline{X}}'$ as follows:

$$g_{a_1,a_2,b}\, x' = (b(x_1 - a_1), \ldots, b(x_{m'} - a_1), b(y_1 - a_2), \ldots, b(y_{n'} - a_2)). \tag{50}$$

For $\Omega = \{(a_1,a_2,b)\}$, $-\infty < a_1,a_2 < \infty$, $0 < b < \infty$, define $g^*_{a_1,a_2,b}$ as follows:

$$g^*_{a_1,a_2,b}(\theta_1,\theta_2,\sigma) = (b(\theta_1 - a_1), b(\theta_2 - a_2), b\sigma). \tag{51}$$

Theorem 4.4.

Let $a_0 \doteq 1,0$ according as $x_2 \geq x_1$ or $x_2 < x_1$, $a_i = \dfrac{x_{i+2} - x_1}{x_2 - x_1}$, $(i = 1, \ldots, m-2)$, $b_j = \dfrac{y_{j+2} - y_1}{y_2 - y_1}$, $(j = 1, \ldots, n-2)$, and

$c = \dfrac{y_2 - y_1}{x_2 - x_1}$. Let E_R denote the conditional expectation over a region R in which the ancillaries a_0, a_i, $(i = 1,\ldots,n-2)$, b_j, $(j=1,\ldots,n-2)$ and c are fixed and E_f denote the expectation with respect to the fiducial density of $\theta_1,\theta_2,\sigma,x^*y^*$ given x,y.

Assume:

(1) $(x',y') = ((x,x^*),(y,y^*))$ has density with respect to the Lebesgue

measure $L_{m'+n'}$ given by:

$$p(x',y',\theta_1,\theta_2,\sigma) = \sigma^{-(m'+n')} f(\frac{x_1-\theta_1}{\sigma},\ldots,\frac{x_{m'}-\theta_1}{\sigma}, \frac{y_1-\theta_2}{\sigma},\ldots,\frac{y_{n'}-\theta_2}{\sigma}), \tag{52}$$

$-\infty < x_1,\ldots,x_{m'}, y_1,\ldots,y_{n'} < \infty$, where $x = (x_1,\ldots,x_m)$, $x^* = (x_{m+1},\ldots,x_{m'})$, $y = (y_1,\ldots,y_n)$, $y^* = (y_{n+1},\ldots,y_{n'})$, $m' > m$ and $n' > n$.

(2) $H(x',y,\theta_1,\theta_2,\sigma)$ is an invariant function i.e. for which (10) holds. Then:

$$E_R(H(x',y',\theta_1,\theta_2,\sigma)) = E_f(H(x',y',\theta_1,\theta_2,\sigma)). \tag{53}$$

Proof:

The proof of this theorem will not be given. It is similar to the proof of Theorems 4.2. and 4.3.

Definition 4.13.

A function $B(x_1, \ldots, x_m, y_1, \ldots, y_n)$ will be said to be an invariant predictor if for all a_1,a_2,b, $-\infty < a_1,a_2 < \infty$ and $b > 0$, we have that:

$$b[B(x_1, \ldots, x_m, y_1, \ldots, y_n) - (a_1 - a_2)] = B[b(x_1 - a_1), \ldots, b(x_m - a_1), b(y_1 - a_2), \ldots, b(y_n - a_2)]. \tag{54}$$

Lemma 4.5.

Let $H(x',y',\theta_1,\theta_2,\sigma) = \varphi[\frac{1}{\sigma}(B(x_1,\ldots,x_m,y_1,\ldots,y_n) - (x_{m+1} - y_{n+1}))]$.

Then (10) holds for $H(x',y',\theta_1,\theta_2,\sigma)$.

Proof:

The proof of this is immediate and will not be given.

Corollary 4.6.

If Assumption (1) of Theorem 4.4. is satisfied, $x = (x_1, \ldots, x_m)$, $x^* = x_{m+1}$, $y = (y_1, \ldots, y_n)$ and $y^* = y_{n+1}$, then the minimum mean square error invariant predictor of $x_{m+1} - y_{n+1}$ is $E_f(\frac{x_{m+1} - y_{n+1}}{\sigma^2}) \mid E_f(\frac{1}{\sigma^2})$.

Proof:

If we use Lemma 4.5. and an argument similar to Corollary 4.5., the proof of this Corollary is straight forward.

4.6. Case of Two Location Parameters and Two Scale Parameters.

In this Section we consider a theorem which deals with families of distributions with two location parameters and two scale parameters. For this case we have not been able to obtain or even define in a satisfactory way an invariant predictor for the difference of two future observations i.e. $x_{m+1} - y_{n+1}$.

Definition 4.14.

For $\underline{X}' = R_{m'+n'}$ define g_{a_1,b_1,a_2,b_2} on $\underline{X}'$ as follows:

$$g_{a_1,b_1,a_2,b_2}(x',y') = (b_1(x_1-a_1),\ldots,b_1(x_{m'}-a_1),b_2(y_1-a_2),\ldots,b_2(y_{n'}-a_2)). \quad (55)$$

For $\Omega = \{(a_1,b_1,a_2,b_2)\}, -\infty < a_1,a_2 < \infty,\ 0 < b_1,b_2 < \infty$, define

$$g^*_{a_1,b_1,a_2,b_2}(\theta_1,\sigma_1,\theta_2,\sigma_2) = (b_1(\theta_1-a_1),b_1\sigma_1,b_2(\theta_2-a_2),b_2\sigma_2). (56)$$

Theorem 4.5.

Let $a_0 = 1,0$ according as $x_2 \geq x_1$ or $x_2 < x_1$ and $b_0 = 1,0$ according as $y_2 \geq y_1$ or $y_2 < y_1$. Let $a_i = \dfrac{x_{i+2} - x_1}{x_2 - x_1}$, $(i = 1,\ldots,m-2)$ and $b_j = \dfrac{y_{j+2} - y_1}{y_2 - y_1}$, $(j = 1,\ldots,n-2)$. Also let E_R denote the conditional expectation over a region R in which the ancillaries $a_0,a_1,\ldots,a_{m-2}$, $b_0,b_1,\ldots,b_{n-2}$ are fixed and E_f denote the expectation with respect to the fiducial density of $\theta_1,\theta_2,\sigma_1,\sigma_2,x^*$ and y^* given x and y.

Assume that:

(1) $(x',y') = ((x,x^*),(y,y^*))$ has density with respect to the

Lebesgue measure $L_{m'+n'}$ given by:

$$p(x',y',\theta_1,\sigma_1,\theta_2,\sigma_2) = \sigma_1^{-m'} \sigma_2^{-n'} f(\frac{x_1-\theta_1}{\sigma_1},\ldots,\frac{x_{m'}-\theta_1}{\sigma_1},\frac{y_1-\theta_2}{\sigma_2},\ldots,\frac{y_{n'}-\theta_2}{\sigma_2}), \tag{57}$$

$-\infty < x_1,\ldots,x_{n'}, y_1,\ldots,y_{m'} < \infty$, where $x = (x_1,\ldots,x_n)$, $x^* = (x_{n+1},\ldots,x_{n'})$, $m' > m$ and $n' > n$.

(2) $H(x',y',\theta_1,\theta_2,\sigma_1,\sigma_2)$ is an invariant function i.e. for which (10) holds. Then:

$$E_R(H(x',y',\theta_1,\theta_2,\sigma_1,\sigma_2)) = E_f(H(x',y',\theta_1,\theta_2,\sigma_1,\sigma_2)). \tag{58}$$

Proof:

The proof of this theorem is similar to the proof of Theorem 4.2. and 4.3. and so will not be given.

Definition 4.15.

$B(x_1,\ldots,x_m,y_1,\ldots,y_n)$ will be said to be an invariant ratio predictor if $B(x_1,\ldots,x_m,y_1,\ldots,y_n) = \dfrac{B_1(x_1,\ldots,x_m)}{B_2(y_1,\ldots,y_n)}$ and $B_1(x_1,\ldots,x_m)$ and $B_2(y_1,\ldots,y_n)$ are invariant functions in the sense of (47).

Lemma 4.6.

Let $S(x_{m+1},\ldots,x_{m'},y_{n+1},\ldots,y_{n'})$, $m' > m$ and $n' > n$ be any function of the future observations which has the property of invariance according to Definition 4.15. Then:

$\varphi(\frac{\sigma_2}{\sigma_1}[B(x_1,\ldots,x_m,y_1,\ldots,y_n) - S(x_{m+1},\ldots,x_{m'},y_{n+1},\ldots,y_{n'})])$ satisfies (10).

Proof:

The proof of this is straight forward and will not be given.

Corollary 4.7.

If assumption(1) of Theorem 4.5. is satisfied and $S(x_{m+1},\ldots,x_{m'},y_{n+1},\ldots,y_{n'})$, $m' > m$, $n' > n$, is a function of the future observations which has the property of invariance according to Definition 4.15., then $E_f(\frac{S\sigma_2^2}{\sigma_1^2}) \,/\, E_f(\frac{\sigma_2^2}{\sigma_1^2})$ is a minimum mean square error invariant predictor of S.

Proof:

If we use Lemma 4.6. and an agrument similar to Corollary 4.5., the proof of this Corollary is straight forward.

4.7. Some Remarks:

Several other cases have not been considered, e.g. the cases of (1) two location parameters and (2) k location parameters and k scale parameters $(k > 2)$. We want to mention here that the cases are also special cases of Theorem 4.1. However, as remarked in Section 4.6., we may not always be able to obtain suitable invariant predictors.

Chapter 5. Invariant Functions.

5.0. Introduction.

In this Chapter we consider invariant functions on the sample and parameter spaces i.e. functions for which (11) of Chapter 2 holds. Lehmann (1959) has dealt with invariant functions on the sample space. He has obtained theorems concerning: (i) the relationship of the maximal invariant with an invariant function, (ii) a method for obtaining a maximal invariant, (iii) the manner in which the parameter space can be shrunk by use of maximal invariant on the parameter space. Theorems corresponding to (i) and (ii) above have been obtained. The definition of maximal invariant which we give below is an obvious extension of Lehmann's definition of maximal invariant on the sample space. We begin with assumptions which are similar to those of Chapter 2, but fewer in number and in some respects similar.

Assumption 1.

Let $(\underline{\overline{X}}, \beta_{\underline{\overline{X}}}, P^{\omega})$ be a probability space, where $\omega\epsilon\Omega$ and Ω is the parameter space.

Assumption 2.

There is a group $\mathcal{G} = \{g\}$ of (1-1) measurable transformations on the sample space $\underline{\overline{X}}$ onto itself.

Assumption 3.

The class of measures P^{ω} for $\omega\epsilon\Omega$ is closed under $\mathcal{G}$. Thus for any $g\epsilon \mathcal{G}$ and $\omega\epsilon\Omega$, there is a $\omega_g\epsilon\Omega$ such that for all $X\epsilon\beta_{\underline{\overline{X}}}$,

$$P^{\omega}(X) = P^{g^{*}\omega}(gX), \qquad (1)$$

where g^{*} is the function on Ω to Ω defined by $g^{*}\omega = \omega_g$.

The transformations $\mathcal{G}^{*} = \{g^{*}\}$ is a group.

<u>Assumption 4.</u>

For any $\omega_1, \omega_2 \in \Omega$, there is a single $g \in \mathcal{G}$ such that $g^{*}\omega_1 = \omega_2$.

5.1. Relation Between a Maximal Invariant and Invariant Function.

In this Section we derive a relation between a maximal invariant and an invariant function on the sample and parameter spaces.

Definition 5.1.

A function $S(x,\omega)$ will be said to be maximal invariant under $\mathcal{G}, \mathcal{G}^*(\mathcal{G} \longleftrightarrow \mathcal{G}^*)$ if:

(1) $S(x,\omega)$ is invariant i.e. (11) of Chapter 2 holds.

(2) $S(x_1,\omega_1) = S(x_2,\omega_2)$ implies that there exists $g \in \mathcal{G}$ and $g^* \in \mathcal{G}^*$ $(g \longleftrightarrow g^*)$ such that $gx_1 = x_2$ and $g^*\omega_1 = \omega_2$.

Theorem 5.1.

Assume:

(1) Assumptions (1) - (4) hold.

(2) $S(x,\omega)$ is a miximal invariant under $\mathcal{G}, \mathcal{G}^*$ $(\mathcal{G} \longleftrightarrow \mathcal{G}^*)$.

Then a necessary and sufficient condition for $H(x,\omega)$ to be invariant is that it depends on x,ω only through $S(x,\omega)$; i.e. there exists a function h such that $H(x,\omega) = h(S(x,\omega))$, for all x,ω.

Proof:

If $H(x,\omega) = h(S(x,\omega))$, for all x,ω, then $H(gx,g^*\omega) = h[S(gx,g^*\omega)]$ $= H(x,\omega)$, by hypothesis (2) of the theorem. Conversely, if $H(x,\omega)$ is invariant, and if $S(x_1,\omega_1) = S(x_2,\omega_2)$, then for some $g \in \mathcal{G}$, $g^* \in \mathcal{G}^*(g \longleftrightarrow g^*)$, $gx_1 = x_2$ and $g^*\omega_1 = \omega_2$ and therefore, $H(x_1,\omega_1) = H(x_2,\omega_2)$. This completes the proof of the theorem.

Below we give examples of functions which are maximal invariant in the sense of Definition 5.1. and also of functions which are invariant in the sense (11) of Chapter 2 but not maximal invariant.

Example 5.1.

Example of a Maximal Invariant in Case of One Location Parameter Families.

Let:

(1) (R_n, B_n, P^θ) be a probability space, where P^θ, $-\infty < \theta < \infty$, is a class of measures.

(2) $\mathcal{G} = \{g_\theta : -\infty < \theta < \infty\}$ be a group of transformations on R_n, where g_θ is defined as follows:

$$g_\theta x = (x_1 - \theta, \ldots, x_n - \theta). \qquad (2)$$

(3) $\mathcal{G}^* = \{g_\theta^* : -\infty < \theta < \infty\}$, be the corresponding group of transformations on the parameter space Ω, where g_θ^* is defined as follows:

$$g_\theta^* (\theta_1) = \theta_1 - \theta. \qquad (3)$$

Then,

$$S(x,\theta) = (x_1 - \theta, \ldots, x_n - \theta), \qquad (4)$$

is a maximal invariant.

Proof:

It is easily verified that Assumptions (1) - (4) hold and $S(x,\theta) = S(gx, g^*\theta)$, where g <----> g^*. Moreover, suppose that $S(x,\theta_1) = S(y,\theta_2)$ i.e. $x_i - \theta_1 = y_i - \theta_2$, $(i = 1, \ldots, n)$. Take $g = g_{x_1 - y_1}$ and hence the corresponding $g^* = g^*_{x_1 - y_1}$. Then it is seen that $g_{x_1 - y_1} x = y$, and $g^*_{x_1 - y_1} \theta_1 = \theta_2$. Thus, $S(g_{x_1 - y_1} x, g^*_{x_1 - y_1} \theta_1) = S(y,\theta_2)$. Hence $S(x,\theta)$ is a maximal invariant.

Example 5.2.

Example of a Maximal Invariant in Case of One Location Parameter and One Scale Parameter Families.

Let:

(1) $(R_n, B_n, P^{\theta,\sigma})$ be a probability space where $P^{\theta,\sigma}$, $-\infty < \theta < \infty$, $\sigma > 0$ is a class of measures.

(2) $\mathcal{G} = \{g_{a,b}: -\infty < a < \infty,\ b > 0\}$ be a group of transformations on R_n, where $g_{a,b}$ is defined as follows:

$$g_{a,b}\, x = (b(x_1 - a), \ldots, b(x_n - a)). \tag{5}$$

(3) Also, $\mathcal{G}^* = \{g^*_{a,b} : -\infty < a < \infty,\ b > 0\}$ be the corresponding group of transformations on the parameter space Ω, where $g^*_{a,b}$ is defined as follows:

$$g^*_{a,b}\,(\theta,\sigma) = (b(\theta - a), b\sigma). \tag{6}$$

Then,

$$S(x,\theta,\sigma) = \left(\frac{x_1 - \theta}{\sigma}, \ldots, \frac{x_n - \theta}{\sigma}\right) \tag{7}$$

is a miximal invariant.

Proof:

It is easily verified that Assumptions (1) - (4) hold and $S(x,\theta,\sigma) = S(gx, g^*(\theta,\sigma))$, where $g \longleftrightarrow g^*$. Moreover, suppose that $S(x,\theta_1,\sigma_1) = S(y,\theta_2,\sigma_2)$ i.e.,

$\frac{x_i - \theta_1}{\sigma_1} = \frac{y_i - \theta_2}{\sigma_2}$, $(i = 1, \ldots, n)$. Take $g = g_{a,b}$ and hence the corresponding $g^* = g^*_{a,b}$, where $a = \frac{y_1 - x_2}{y_1 - y_2}$ and $b = \frac{y_1 - y_2}{x_2 - x_1}$. It is easily seen that we have the following hold:

$$\text{(1)} \quad \frac{y_1x_2 - y_2x_1}{y_1 - y_2} = \theta_1 - \frac{\sigma_1}{\sigma_2}\theta_2. \tag{8}$$

$$\text{(2)} \quad \frac{y_1 - y_2}{x_1 - x_2} = \frac{\sigma_2}{\sigma_1}. \tag{9}$$

By use of (8) and (9) and some routine algebra, we have that $g_{a,b}\,x = y$ and $g^*_{a,b}(\theta_1,\sigma_1) = (\theta_2,\sigma_2)$. Thus $S(x, \ ,\sigma)$ is a maximal invariant.

<u>Example 5.3.</u>

<u>Example of a Function which is Invariant but not Maximal Invariant in Case of One Location Parameter Families.</u>

Let $\varphi(x,\theta) = (x_1-\theta)^2 + (x_2 - \theta)^2$. Take $x = (4,5)$ and $y = (0,5)$, $\theta_1 = 1$ and $\theta_2 = 0$. Then $\varphi(x,\theta_1) = \varphi(y,\theta_2) = 25$. There is no $g\in \mathcal{G}$ and $g^*\in \mathcal{G}^*$, (g <----> g^*) such that $gx = y$ and $g^*\theta_1 = \theta_2$. Thus $\varphi(x,\theta)$ is not maximal invariant, however it can be easily verified that $\varphi(x,\theta)$ is an invariant function.

5.2. A Method for Obtaining Maximal Invariant.

In this Section a method to obtain a maximal invariant through subgroups is suggested. It is an obvious extension of the method by Lehmann (1959) for obtaining maximal invariant on the sample space. We start by giving two lemmas which will be needed later.

Lemma 5.1.

Let Assumptions (1) - (4) hold. Let G be a subgroup of $\mathcal{G}$. Then,

$$G^* = \{g^* : g^* \longleftrightarrow g \in G\} \tag{10}$$

is a subgroup of $\mathcal{G}^*$.

Proof:

The proof of this lemma is similar to Lemma 1 on page 214 of Lehmann (1959).

Lemma 5.2.

(1) Given $g \in \mathcal{G}$, $g^* \in \mathcal{G}^*$, where $g \longleftrightarrow g^*$, let g^{**} be defined on $S(x,\omega)$ as follows:

$$g^{**}(s(x,\omega)) = S(gx, g^*\omega). \tag{11}$$

(2) Let $\mathcal{G}^{**} = \{g^{**} : g^{**} \longleftrightarrow g \in \mathcal{G},\ g^* \in \mathcal{G}^*, \text{ where } g \longleftrightarrow g^*\}$.

(3) Let $G^{**} = \{g^{**} : g^{**} \longleftrightarrow g \in G, g^* \in G^*, \text{ where } g \longleftrightarrow g^*\}$, (12)

where G is a subgroup of $\mathcal{G}$ and G^* is the corresponding subgroup of $\mathcal{G}^*$ according to (10).

Then $\mathcal{G}^{**}$ is a group and G^{**} is a subgroup of $\mathcal{G}^{**}$.

Proof:

We first show that $\mathcal{G}^{**}$ is a group. We have the following:

(1) The identity element is clearly contained in $\mathcal{G}^{**}$.

(2) Suppose g_1^{**} and g_2^{**} are contained in $\mathcal{G}^{**}$. Then, $g_1^{**}(S(x,\omega)) = S(gx_1, g^*\omega)$ and $g_2^{**}(g_1^{**} S(x,\omega)) = S(g_2 g_1 x, g_2^* g_1^* \omega) = (g_2 g_1)^* (S(x,\omega))$. Therefore $g_2^{**} \cdot g_1^{**}$ is contained in $\mathcal{G}^{**}$.

(3) $(g^{-1})^{**}(g^{**}S(x,\omega)) = (g^{-1})^{**} (S(gx,g^*\omega)) = S(x,\omega)$, for $g^{**} \epsilon \mathcal{G}^{**}$. Therefore, $(g^{-1})^{**} = (g^{**})^{-1}$ and $(g^{**})^{-1}$ is contained in $\mathcal{G}^{**}$. Hence, $\mathcal{G}^{**}$ is a group. Similarly it can be shown that G^{**} is a subgroup of $\mathcal{G}^{**}$.

The following theorem is concerned with the obtaining of a maximal invariant through subgroups.

Theorem 5.2.

Let:

(1) D and E be two subgroups generating $\mathcal{G}$.

(2) $S(x,\omega)$ be a maximal invariant with respect to D, D^* ($D \longleftrightarrow D^*$).

(3) $t(S(x,\omega))$ be a maximal invariant under the group E^{**} of transformations, where $E^{**} = \{e^{**} : e^{**} \longleftrightarrow e \epsilon E,\ e^* \epsilon E^*, \text{ where } e \longleftrightarrow e^*\}$.

Assume that:

(4) Assumptions (1) - (4) hold.

(5) For $e \epsilon E$, $e^* \epsilon E^*$, ($e \longleftrightarrow e^*$),

$$S(x_1,\omega_1) = S(x_2,\omega_2) \text{ implies that } S(ex_1, e^*\omega_1) = S(ex_2, e^*\omega_2), \tag{13}$$

$e \longleftrightarrow e^*$. Then, $t(S(x,\omega))$ is a maximal invariant with respect to $\mathcal{G}$, $\mathcal{G}^*$ ($\mathcal{G} \longleftrightarrow \mathcal{G}^*$).

Proof:

Let $S(x_1,\omega_1) = S(x_2,\omega_2)$. Then $t(S(x_1,\omega_1)) = t(e^{**}S(x_1,\omega_1)) =$

$t(e^{**}S(x_2,\omega_2)) = t(S(x_2,\omega_2))$, where the equalities follow from (3), (5) and (3) of the theorem. Hence, $t(\cdot)$ is well defined.

Any element $g \in \mathcal{G}$ has the form $g = e_m d_m \ldots e_1 d_1$. Let $(y_1,\omega_2) = (gx_1, g^*\omega_1)$.

Then $t(S(y_1,\omega_2)) = t(S(gx_1,g^*\omega_1)) = t(S(e_m d_m \ldots e_1 d_1 x_1, e_m^* d_m^* \ldots e_1^* d_1^* \omega_1)$

$= t(e_m^{**} S(d_m e_{m-1} \ldots e_1 d_1 x_1, d_m^* e_{m-1}^* \ldots e_1^* d_1^* \omega_1)) = t(S(d_m e_{m-1} \ldots e_1 d_1 x_1,$

$d_m^* e_{m-1}^* \ldots e_1^* d_1^* \omega_1)) = t(S(e_{m-1} \ldots e_1 d_1 x_1, e_{m-1}^* \ldots e_1^* d_1^* \omega_1))$, where the last three equalities follow from the definition of e_m^{**}, invariance of t with respect to E^{**} and invariance of S with respect to D,D^* $(D \longleftrightarrow D^*)$ respectively. Then, by induction we have that $t(S(y_1,\omega_2)) = t(S(x_1,\omega_1))$ and $t(S(x,\omega))$ is invariant with respect to $\mathcal{G}, \mathcal{G}^*(\mathcal{G} \longleftrightarrow \mathcal{G}^*)$.

Since $t(S(y_1,\omega_2)) = t(S(x_1,\omega_1))$ and $t(S(x,\omega))$ is maximal invariant with respect to E^{**}, we have that there exists a e^{**} such that:

$$S(y_1,\omega_2) = e^{**}(S(x_1,\omega_1)) = S(ex_1,e^*\omega_1). \tag{14}$$

Also, since $S(x,\omega)$ is maximal invariant with respect to $D,D^*(D \longleftrightarrow D^*)$, we have that there exists a $d \in D$, $d^* \in D^*(d \longleftrightarrow d^*)$ such that $y_1 = dex_1$ and $\omega_2 = d^*e^*\omega_1$. But $de \in \mathcal{G}$ and $d^*e^* \in \mathcal{G}^*$. Hence $t(S(x,\omega))$ is a maximal invariant with respect to $\mathcal{G}$, $\mathcal{G}^*(\mathcal{G} \longleftrightarrow \mathcal{G}^*)$.

<u>Example 5.4.</u>

<u>Obtaining of Maximal Invariant in Case of One Location Parameter and One Scale Parameter.</u>

Consider again the problem with specification of Example 5.2. Let $D = \{g_{a,1} : -\infty < a < \infty\}$ and $E = \{g_{0,b} : b > 0\}$, where $g_{a,1}$ and $g_{0,b}$ are defined according to (5). It is easily seen that D and E together generate $\mathcal{G}$. Moreover, D^* and E^*, the corresponding subgroups of $\mathcal{G}^*$ are given by:

$$D^* = \{g^*_{a,1} : -\infty < a < \infty\}, \text{ and } G^* = \{g^*_{0,b} : b > 0\}, \tag{15}$$

where $g^*_{a,1}$ and $g^*_{0,b}$ are defined according to (6). As remarked in Example 5.2., Assumptions (1) - (4) are satisfied. Let,

$$S(x,\theta,\sigma) = (x_1 - \theta, \ldots, x_n - \theta). \tag{16}$$

Then it can be easily shown (in a manner similar to Example 5.1.) that $S(x,\theta,\sigma)$ is a maximal invariant with respect to $D, D^* (D <----> D^*)$. Also, if $(x_1 - \theta, \ldots, x_n - \theta) = (y_1 - \theta_2, \ldots, y_n - \theta_2)$, then

$(b(x_1 - \theta_1), \ldots, b(x_n - \theta_1)) = (b(y_1 - \theta_2), \ldots, b(y_n - \theta_2))$.

Hence (5) of the theorem is satisfied. Finally, define $t(S(x,\theta,\sigma))$ as follows:

$$t(S(x,\theta,\sigma)) = \left(\frac{x_1 - \theta}{\sigma}, \ldots, \frac{x_n - \theta}{\sigma}\right). \tag{17}$$

Let $$\left(\frac{x_1 - \theta_1}{\sigma_1}, \ldots, \frac{x_n - \theta_1}{\sigma_1}\right) = \left(\frac{y_1 - \theta_2}{\sigma_2}, \ldots, \frac{y_n - \theta_2}{\sigma_2}\right) \tag{18}$$

and $e^{**}_{a,b} <----> (e_{a,b}, e^*_{a,b})$, where $(a,b) = (0, \frac{\sigma_2}{\sigma_1})$. Then

$e^{**}_{a,b}(x_1 - \theta_1, \ldots, x_n - \theta_1) = \left(\frac{\sigma_2}{\sigma_1}(x_1 - \theta_1), \ldots, \frac{\sigma_2}{\sigma_1}(x_n - \theta_1)\right) =$

$(y_1 - \theta_2, \ldots, y_n - \theta_2)$, where the last equality follows from (18). Thus all the hypotheses of the theorem are satisfied and (17) is a maximal invariant with respect to $\mathcal{G}, \mathcal{G}^* (\mathcal{G} <----> \mathcal{G}^*)$.

REFERENCES

Breitenberger, Ernst. (1963). "Analogues of the normal distribution on the circle and the sphere." Biometrika, 50, 81-89.

Brillinger, David R. (1962). "Examples bearing on the definition of fiducial probability with a bibliography." Annals of Mathematical Statistics 33, 1349-1355.

Buehler, R. J. (1963). "A new consistency test for fiducial distributions and some related fiducial theory." Technical Report Number 1, National Science Foundation Grant G14831. Statistical Laboratory, Iowa State University, Ames.

Fisher, R. A. (1935). "The fiducial argument in statistical inference." Annals of Eugenics 6, 391-398.

Fisher, R. A. (1956). Statistical Methods and Scientific Inference. Edinburgh; Oliver and Boyd.

Fraser, D. A. S. (1961). "The fiducial method and invariance." Biometrika 48, 261-280.

Halmos, P. R. (1950). Measure Theory. New York; Van Nostrand.

Lehmann, E. L. (1959). Testing Statistical Hypotheses. New York; Wiley.

Owen, A. R. G. (1948). "Ancillary statistics and fiducial distributions." Sankhya 9, 1-18.

Pitman, E. J. G. (1939). "The estimation of the location and scale parameters of a continuous population of any given form." Biometrika 30, 391-421.

Ramsey, F. L. and Buehler, R. J. (1963). "Prediction in location and scale parameter families." Technical Report Number 2, National Science Foundation Grant G14831. Statistical Laboratory, Iowa State University, Ames.

Segal, I. E. (1938). "Fiducial Distribution of several parameters with application to a normal system." Proceedings of the Cambridge Philosophical Society 34, 41-47.

Tukey, John W. (1957). "Some examples with fiducial relevance." Annals of Mathematical Statistics 28, 687-695.

ACKNOWLEDGEMENTS

I am grateful to Professor R. J. Buehler who introduced me so rapidly to this area of statistics. His inspiring guidance and patience for having conversations with me provided tremendous incentive. My thanks are also due to Mrs. Harold Criswell and Mrs. Daniel Johnson for typing the manuscript. Acknowledgement here is also made to the National Science Foundation for financial support under Grant G19126.